

The metal–insulator transition: a perspective

BY P. P. EDWARDS¹, R. L. JOHNSTON¹, C. N. R. RAO², D. P. TUNSTALL³
AND F. HENSEL⁴

¹*School of Chemistry, University of Birmingham,
Edgbaston, Birmingham B15 2TT, UK*

²*CSIR Centre of Excellence in Chemistry and Solid State and Structural Chemistry
Unit, Indian Institute of Science, Bangalore 560012, India*

³*School of Physics and Astronomy, University of St Andrews,
St Andrews KY16 9AJ, UK*

⁴*Institute of Physical Chemistry, Nuclear Chemistry and Macromolecular
Chemistry, University of Marburg, Hans-Meerwein-Strasse,
D.35032 Marburg, Germany*

The metal–insulator transition, a quantum phase transition signifying the natural transformation of a metallic conductor to an insulator, continues to be the focus of intense inquiry and debate. The first discussion of the heuristic differences between metals and insulators, and implicitly the critical conditions for the transition between these canonical electronic regimes, dates back to the dawn of the twentieth century. As we approach the end of the century, the precise nature of the metal–insulator transition remains one of the major intellectual challenges in condensed matter science. In this article we present a brief introduction to just some of the key underlying features of this enduring physical phenomenon. The following articles and discussion present a detailed current account of the many facets of the science of the metal–insulator transition.

Keywords: metal–non-metal transition; metal–insulator transition; minimum metallic conductivity; electrical conductivity; electron–electron interactions

1. The metallic and non-metallic states of matter

The electrical conductivity of solids ranges from at least $10^9 \Omega^{-1} \text{cm}^{-1}$ for a pure metal such as copper at liquid helium temperatures to at most $10^{-22} \Omega^{-1} \text{cm}^{-1}$ for the best insulators or non-metals at this same base temperature. Bardeen (1940) and later McMillan (1963) first drew attention to these vast differences in electrical conductivity separating the metallic and non-metallic states of matter, and commented that this variation, amounting to a factor of at least 10^{31} , probably represents the widest variation for any (laboratory-measurable) physical property (see also Ehrenreich 1967).

The fundamental issue as to precisely *why* certain materials are excellent conductors, while others patently are not, has a long and venerable history (Mott 1985) dating back to the dawn of the 20th century. More recently, Feynman *et al.* (1964) alluded to these same issues in his celebrated Lectures on Physics, noting ‘some materials are electrical “conductors”, because their electrons are free to move about;

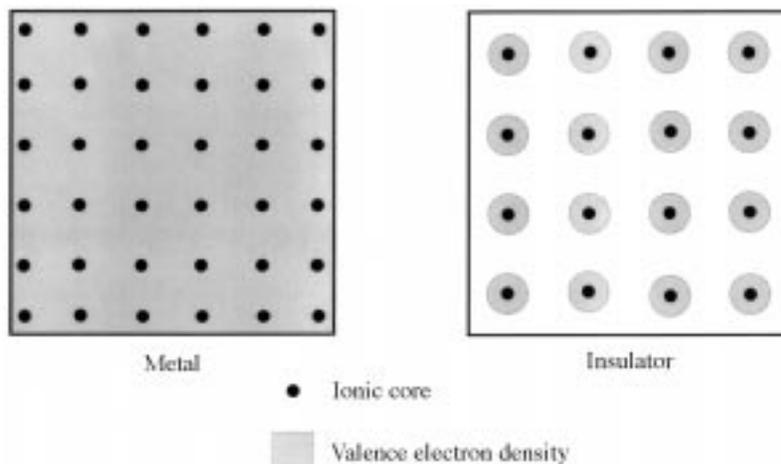


Figure 1. An artist's impression of the metal–insulator transition at zero temperature, whereby itinerant (delocalized) electrons become localized at individual sites; here the system transforms from a metal ($\sigma(T = 0 \text{ K}) \neq 0$) to an insulator ($\sigma(T = 0 \text{ K}) = 0$).

others are insulators, because their electrons are held tightly to individual atoms. We shall consider later how some of these properties come about, but that is a very complicated subject.' If the existence of metals versus insulators is indeed a 'complicated subject', the issue of how each phase can transform into the other, the metal–insulator transition, is equally daunting. And yet, remarkably, there are countless examples in nature where metals, those magnificent conductors of electricity (Cottrell 1992), can be continuously transformed into stubbornly resistive insulators, and vice versa (Mott 1974, 1990; Edwards & Rao 1985, 1995). The metal–insulator transition, the process of physically and chemically transforming a metal into an insulator, and vice versa, has so far proven surprisingly recalcitrant to a complete theoretical analysis (Edwards *et al.* 1995).

At present the subject continues to thrive and develop; it represents a perfect example of a wide-ranging complex and fundamental, but unresolved, scientific question in condensed matter science. For how can we begin to understand the microscopic process, or processes, by which a highly conducting material containing *ca.* 10^{23} free, or itinerant, electrons, all jostling and interacting with one another in the most complicated ways (Cottrell 1997), can suddenly transform to a situation in which every single one of these itinerant electrons now finds itself localized and completely bound to an individual atomic site in an insulator (figure 1)?

In this short introductory article we will briefly discuss just some of the underlying physical concepts, ideas and models relating to the transformation of a metal to an insulator, or equivalently, an insulator into a metal. Many of the founding insights and major developments originate from the seminal contributions of Sir Nevill Mott to the intriguing problem of the metal–insulator transition (Mott 1937, 1949, 1956, 1958, 1961, 1974, 1985, 1990).

The history of the subject of the metal–insulator transition can be broadly classified into two avenues of development. First, there is the continual effort both to locate and to describe precisely the actual transition point between these two canonical electronic regimes (Mott 1990; Edwards & Rao 1995; Edwards *et al.* 1995). As noted recently by Ramakrishnan (1995), such evolutionary advances of this first kind can often be eclipsed by a second kind of development; namely, major and unexpected

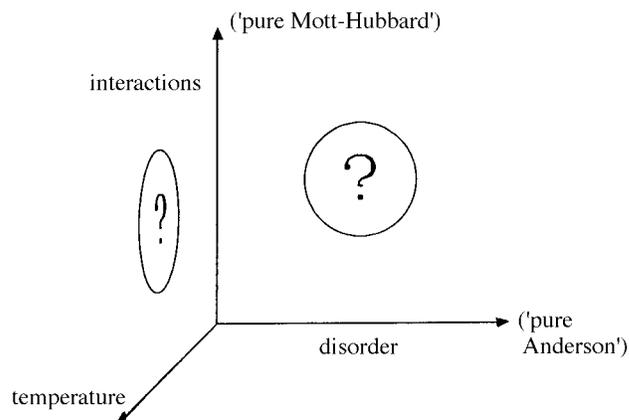


Figure 2. A representation of the intrinsic complexity, and hence the challenge, of the metal–insulator transition; the various competing, and complementary, effects of electron interactions, disorder and temperature. All experimental systems fall somewhere in the three-dimensional space, away from the origin, which represents the zero-temperature non-disordered non-interacting situation. Taken from Logan *et al.* (1995).

experimental discoveries which initiate and dictate a reassessment and expansion of the problem of the metal–insulator transition. Recent examples would be the discovery of high temperature superconductivity in cuprates close to the metal–insulator transition (for a review, see Iye 1995; Rao 1996; Edwards *et al.* 1995, 1998) and the emergence of giant magnetoresistive effects in marginally metallic transition metal oxides (for a review, see Rao 1996). This healthy synergy between theory and experiment continues to be a characteristic hallmark of the subject of the metal–insulator transition as we review it here.

2. How can we visualize the transition?

The deceptively simple task of describing and understanding the passage from the metallic to the insulating regime (figure 1) still remains an enigma. Of course, part of the problem and indeed the fascination is that any metal–insulator transition almost certainly does not occur by a single mechanism but, instead, may arise from a variety of competing, but complementary, electronic mechanisms. These involve the close interplay of various contributing features, for example, disorder, electron–electron interactions, screening, etc. (Mott 1974, 1990). Throw into this complex problem the crucial role of temperature and the situation becomes even more interesting, and even more taxing! The scale of the problem at hand is hinted at in figure 2, which is a representation of three important facets of the problem, namely temperature, electron–electron interactions and disorder (Logan *et al.* 1995). Here, the reader is reminded that most of conventional band theory is confined to a single point at the origin, the non-interacting non-disordered limit, typically representative of a situation only at $T = 0$ K. We note also that *all* experimental systems lie away from the origin (figure 2). In these regions the mechanism of electron localization versus itinerancy depends critically upon the interplay and complementarity of these contributing features (Mott 1974, 1990; Logan *et al.* 1995).

Interestingly, the more the intellectual pursuit moves to the consideration of the actual transition point from metal to insulator (and vice versa), the greater the degree

of (apparent) sophistication and complexity necessary for any theoretical description. Paradoxically, it also transpires that simple, but highly effective criteria such as the Mott metallization condition, the Mott minimum metallic conductivity, the Herzfeld polarization catastrophe criterion, etc., have continually emerged as powerful paradigms for describing key features of the metal–insulator transition (Edwards 1995; Edwards *et al.* 1995). It must also be noted here that even the ‘simplest’ of such theoretical models contain a surprising richness and subtlety that perhaps only now is becoming apparent (Ashcroft 1993; Logan & Edwards 1985).

(a) *Polarization, ionization and screening*

Probably the first quantitative attempt to explain the occurrence of metallic versus insulating behaviour in a material, and with it the first discussion of the metal–insulator transition, was made by Goldhammer (1913) and Herzfeld (1927). This elegant rationalization, in terms of relevant atomic properties, which in a sense confer metallic versus insulating status upon any element or material, leads to what is commonly called the Goldhammer–Herzfeld criterion for metallization.

It is important to note that the Goldhammer–Herzfeld view of the metal–insulator transition predates any quantum-mechanical description of the phenomenon, and is based on the density-induced changes to the electronic polarizability (α) of a free atom brought about by the presence of electric fields generated by neighbouring, and distant, atoms within a condensed phase. With increasing elemental density, a critical divergence in the electronic polarizability (and hence in the dielectric constant) is predicted, causing the catastrophic release (or wholesale freeing) of *all* bound valence electrons, with concomitant metallization and high electrical conductivity.

The Goldhammer–Herzfeld view can most profitably be viewed in terms of the Clausius–Mossotti relationship (Herzfeld 1927; Edwards & Sienko 1982, 1983; Logan & Edwards 1985),

$$\frac{(n^2 - 1)}{(n^2 + 2)} = \frac{R}{V}, \quad (2.1)$$

where n is the index of refraction (the high frequency dielectric constant), R is the molar refractivity ($\frac{4}{3}\pi N\alpha$), N is the Avogadro number and V is the molar volume. Herzfeld (1927) argued that if we start with a polarizable atom in the gas phase and transform it to the condensed liquid or solid phase, continuously increasing the elemental density such that the ratio (R/V) increases, then for the critical condition, $(R/V) = 1$, we have the equality $(n^2 - 1) = (n^2 + 2)$, i.e. the dielectric constant must now become infinite. This is the so-called polarization or dielectric catastrophe whereby the valence electrons, which before had been quasi-elastically bound to their parent atoms, are now spontaneously ionized and set free via the strong attractive interactions with the multitude of other polarizable atoms in the dense liquid or solid. The concept of the wholesale ‘freeing’ of valence electrons from their parent (atomic) sites to form a metallic conductor has an attractive physical, chemical and conceptual basis (Berggren 1974, 1978; Edwards & Sienko 1982, 1983; Logan & Edwards 1985). Under such critical conditions we see the inability of individual atoms to retain their valence electrons in the face of fierce competition from attractive forces provided by the multitude of other atoms in the condensed phase (figure 1). Clearly, the larger the atomic polarizability and the elemental density, the more intense is this competition. This important link between atomistic (polarizability) and elemental (density) considerations is beautifully captured in Herzfeld’s original

paper (1927), entitled ‘On atomic properties which make an element a metal’. A clear demonstration of the continuing utility of this attractive and powerful descriptor of the metal–insulator transition is given in figure 3, which reveals how the simple Herzfeld criterion can very effectively delineate between metals and insulators in the periodic classification of the elements (Edwards & Sienko 1982, 1983).

This classical viewpoint also allows one readily to estimate the critical conditions required for metallization of any element of the periodic classification, if R is known to a reasonable degree of accuracy and, of course, with V , the molar volume, determined by the elemental density under the experimental conditions in question.

A recent important application of the Herzfeld criterion relates to the first ever metallization of fluid, elemental hydrogen at high pressure (Weir *et al.* 1996; Hensel & Edwards 1996*a, b, c*; Eggen *et al.* 1997). It has long been presumed that hydrogen at sufficiently high density (pressure) would eventually succumb to metallization (Wigner & Huntington 1936). In figure 4, we show the measured electrical conductivity (σ) for compressed fluid hydrogen (Weir *et al.* 1996) together with the corresponding conductivities of the expanded alkali metal fluids rubidium and caesium, all elements measured at comparable temperatures (2000–3000 K) over a wide range of molar (atomic) densities (Hensel & Edwards 1996*b, c*).

It is manifestly obvious that all three of these group 1 elements of the periodic classification undergo a continuous density-induced transition from an insulating to a conducting state. In the case of fluid hydrogen at these high temperatures, pressures close to 2 Mbar are necessary to effect the transition to the conducting state (Weir *et al.* 1996). The ‘conventional’ alkali metals rubidium and caesium, unquestionably metallic at room pressure and temperature, now continuously transform to a state of exceptionally low electrical conductivity by expansion to low elemental density (Freyland & Hensel 1985; Hensel 1996).

The predicted metallization densities derived from the Herzfeld criterion are $0.595 \text{ mol cm}^{-3}$ ($3.59 \times 10^{23} \text{ cm}^{-3}$) for hydrogen, $8.38 \times 10^{-3} \text{ mol cm}^{-3}$ ($5.05 \times 10^{21} \text{ cm}^{-3}$) for rubidium and $6.66 \times 10^{-3} \text{ mol cm}^{-3}$ ($4.01 \times 10^{21} \text{ cm}^{-3}$) for caesium (Hensel & Edwards 1996*b, c*). These estimates (represented by arrows for each element in figure 4) are in excellent agreement with the experimental densities at which the elements hydrogen, rubidium and caesium all attain a limiting value for the electrical conductivity in the region of $2000 \Omega^{-1} \text{ cm}^{-1}$. This value is close to the Ioffe–Regel value for the so-called minimum metallic conductivity (§ 2*c*) of such a high temperature fluid close to the metal–insulator transition (Mott 1974; Hensel & Edwards 1996). The data (figure 4) and such considerations clearly demonstrate the onset of the density-induced metallization in these elements at densities close to the respective Herzfeld estimates. The relatively small value of α for atomic hydrogen (0.67 \AA^3) prescribes the unusually high densities required for the metallization of fluid hydrogen. In contrast, the large polarizabilities for atomic rubidium (47.3 \AA^3) and caesium (59.7 \AA^3) ensure that the heavier members of group 1 are already metallic at elemental densities commensurate with room pressures and temperatures on this planet.

Pauling (1983, personal communication) outlined a simple argument which goes some way to explaining the undoubted success of the Herzfeld criterion (for a more detailed discussion, see Logan & Edwards (1985)). He noted that the cube root of the molar refractivity, R , can be approximated as a characteristic radius of the outermost (valence) electrons in the isolated atom. If this radius is approximately equal to the cube root of the molar (atomic) volume (V), the outer orbitals from one atom will

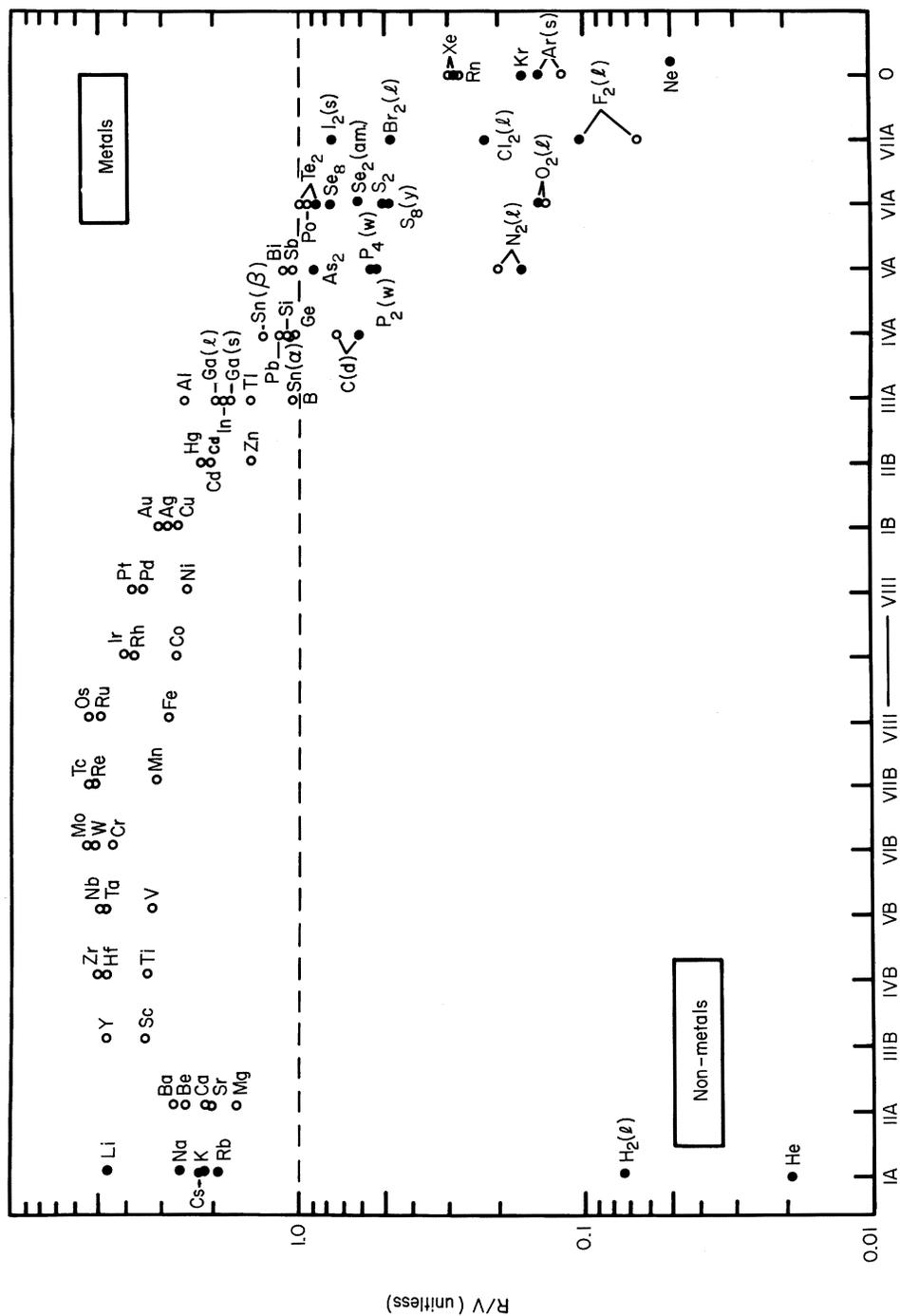


Figure 3. The metallization of elements of the periodic classification under standard temperature and pressure conditions. The figure shows the ratio (R/V) for elements of the periodic classification. The shaded circles represent elements for which R and V are known experimentally. The open circles are for elements for which only V is known experimentally and R is calculated. Taken from Edwards & Sienko (1983).

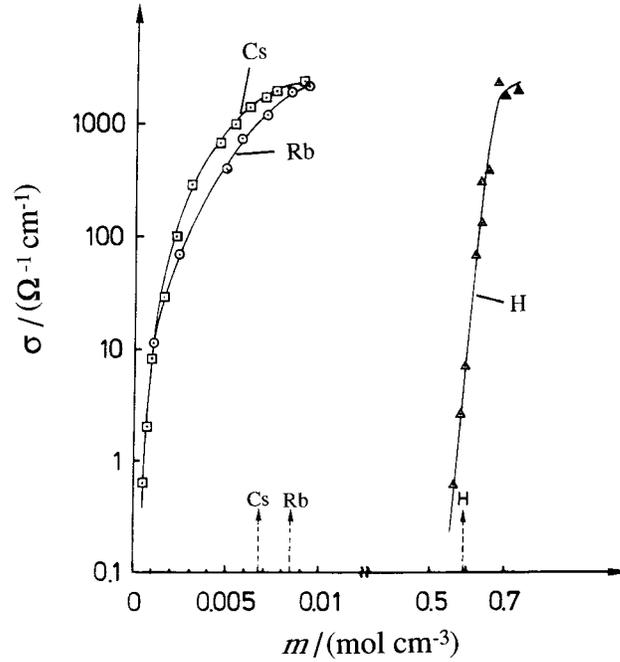


Figure 4. The measured electrical conductivity of fluid caesium, rubidium and hydrogen as a function of the molar atom density m at a temperature of $kT \sim 0.15$ eV. The arrows indicate the predicted metallization densities for each element, based on the Goldhammer–Herzfeld model (see text). Taken from Hensel & Edwards (1996*b, c*).

overlap with those from an adjacent atom and a ‘metallic orbital’ (Pauling 1938) will ensue, with any covalent chemical bonds showing unsynchronized resonance (Pauling 1984), and the element becomes a metallic conductor.

It is also possible to establish a direct link between the Goldhammer–Herzfeld view of the metal–insulator transition and that developed later by Mott (1949, 1956, 1961) in relation to Thomas–Fermi screening and metallization (Edwards & Sienko 1983). Approached from the metallic regime, the metal–insulator transition takes place when the coulombic (attractive) potential ($V(r)$) of an electron–hole pair becomes insufficiently screened via the sea of itinerant conduction electrons and a bound localized state ensues (Mott 1961; Ashcroft 1993). Approached from the insulating regime, this could also be viewed in terms of a polarization or dielectric catastrophe at a critical electron (carrier) density (n_c) when the coulombic potential binding the electron–hole pair drops to zero; electrons are thereby ionized from their constituent atoms or centres (figure 1) and metallization ensues (Edwards & Sienko 1983). Here we have,

$$V(r) = -e^2/\epsilon r, \quad (2.2)$$

where ϵ is the effective dielectric constant of the system, and as $n \rightarrow n_c (= 3/(4\pi\alpha))$, $\epsilon \rightarrow \infty$. Thus the binding energy of the localized electron–hole pair is now reduced to zero at the critical carrier density, n_c . The insulator–metal transition then takes place at n_c as the valence electrons become unbound (ionized) from their parent (hole) sites.

Mott’s theory produced the simple but potent criterion,

$$n_c^{1/3} a_H^* \sim 0.25, \quad (2.3)$$

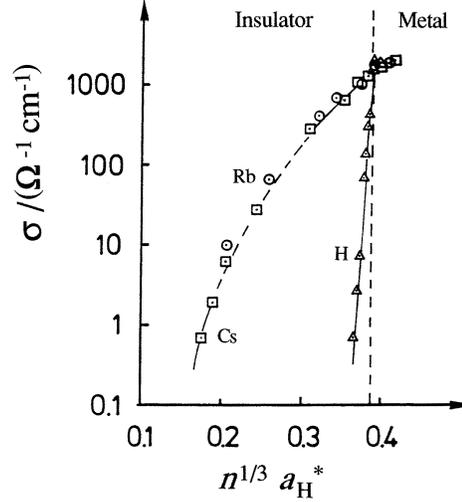


Figure 5. The transition to the metallic state for fluid caesium, rubidium and hydrogen; the dependence of the measured electrical conductivity on the Mott scaling parameter $[n^{1/3} a_H^*]$. The dotted line drawn at $n^{1/3} a_H^* = 0.38$ indicates the common metallization condition for these three alkali elements. Taken from Hensel & Edwards (1996*b, c*).

for the critical conditions at the metal–insulator transition, where n_c is again the critical density of carriers and a_H^* is a characteristic orbital radius of the localized electron centre (Mott 1956, 1961). This venerable criterion, first developed over 35 years ago by Sir Nevill Mott for doped semiconductors, was extended by Edwards & Sienko (1978) to encompass a wide range of experimental systems. The Mott criterion is now known to provide an excellent guide to the location of the metal–insulator transition for over ten orders-of-magnitude of n_c (Edwards *et al.* 1995; Hensel & Edwards 1996*b*). It is also recognized that the critical density for a polarization or dielectric catastrophe (*à la* Herzfeld) is given by $n_c^{1/3} a_H^* \sim 0.38$; and this numerical proximity also establishes a direct link with Mott’s theory (Berggren 1974, 1978; Fritzsche 1978; Edwards & Sienko 1983).

A recent example of the application of the enduring Mott criterion can also be found in the experimental metallization densities of hydrogen, rubidium and caesium (Hensel & Edwards 1996*b, c*). In figure 5 we show the evolution of the density-induced insulator–metal transition in these elements, showing the variation of the measured electrical conductivity with the product $[n^{1/3} a_H^*]$, where n is the electron density and a_H^* is now taken as the radius of the principle maximum in the charge density of the respective valence orbital (e.g. 1s, 5s and 6s, for hydrogen, rubidium and caesium, respectively).

The clear change in the slopes of σ versus $n^{1/3} a_H^*$ at the computed Ioffe–Regel value of the conductivity ($\sigma \sim 2000 \Omega^{-1} \text{cm}^{-1}$) leads one to conclude that all three of these high-temperature fluids become metallic at a constant value of the scaling parameter $n^{1/3} a_H^* \sim 0.38$. Thus, under the appropriate experimental conditions illustrated in figure 5 (namely $n^{1/3} a_H^* > 0.38$) the three fluid elements, hydrogen, rubidium and caesium, can unambiguously be identified as metallic, with hydrogen now assuming its position as the lightest metal in the periodic classification of the elements (Hensel & Edwards 1996*a, b, c*; Eggen *et al.* 1997).

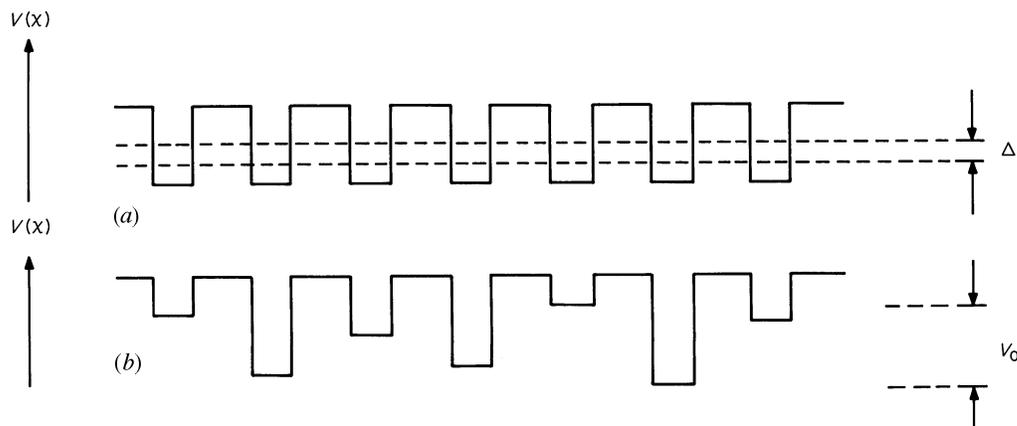


Figure 6. The potential energy of an electron within (a) a periodic field, and (b) a random potential field. Here Δ is the one-electron bandwidth in the absence of the random potential field, V_0 . Adapted from Mott (1974).

(b) *Electron–electron interactions and disorder*

Electronic states involved in charge (electron) transport are spatially extended in a metal, but are localized in an insulator (figure 1). It is now known that electronic localization can be initiated by static disorder (the so-called Anderson transition, (Anderson 1958)), by strong electron–electron interactions or correlations (the so-called Mott–Hubbard transition (Mott 1974; Hubbard 1963, 1964*a, b*)), or by strong electron–lattice coupling (Mott 1974, 1990; Edwards & Rao 1985, 1995). It is also abundantly clear that in all these instances the description of the electronic states in the two limiting regimes, metal and insulator, are qualitatively different. Whilst accurate descriptions have been developed over many years for the extreme (limiting) situations of a metal and an insulator, it is inherently difficult to develop a unified approach which naturally links both electronic regimes across the metal–insulator divide. In addition, for the vast majority of experimental situations, more than one of these individual electronic mechanisms will be operative (figure 2) and each one may reinforce or complement another (Mott 1990; Logan *et al.* 1995; Edwards *et al.* 1995; Ramakrishnan 1995).

Turning first to the idea of localization due to disorder, it was originally pointed out by Anderson (1958) that if the randomness in electronic-state energies at different sites is large enough, electrons, presumed itinerant, become spatially localized. Anderson (1958) used a model of a crystalline array of random potential wells to demonstrate that, provided the disorder-induced potential fluctuations were sufficiently large, an electron could be localized in a finite region of space (figures 6 and 7). Given that for weak levels of disorder, itinerant electrons diffuse, and for strong disorder, they clearly do not, there must be a critical degree of disorder at which an electron at a particular energy becomes localized in space, *vis-à-vis*, the system becomes insulating, such that the DC electrical conductivity tends to zero as the temperature approaches absolute zero ($\sigma(T = 0 \text{ K}) \rightarrow 0$). The corresponding definition of a metal would therefore be $\sigma(T = 0 \text{ K}) \neq 0$, representing a finite value of the electrical conductivity at the absolute zero of temperature.

Thus, as the degree of disorder increases, a metallic conductor can continuously transform into an insulator (Anderson 1958; Mott 1974). The critical energy separating the band of metallic extended states from those of the localized states charac-

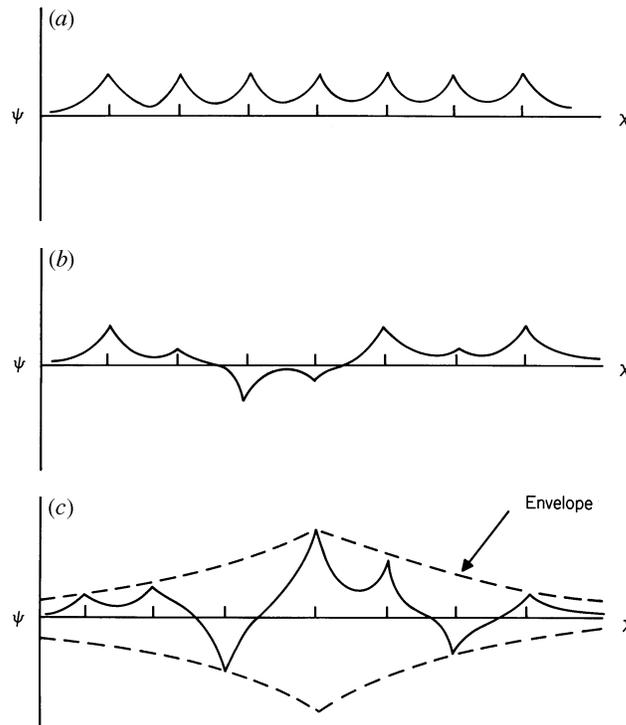


Figure 7. A typical wavefunction ψ in an Anderson lattice: (a) an extended wavefunction of Bloch character, with an electronic mean free path far exceeding the separation between potential wells, and $\Delta \gg V_0$; (b) an extended wavefunction subject to a disorder potential and $\Delta > V_0$; (c) a weakly localized function for which $\Delta \sim V_0$. The overall form of the envelope function is sketched in (c) for weak localization. Adapted from Mott (1974).

teristic of the insulator has been termed the mobility edge (Mott 1974). As disorder and/or the electron density changes, the Fermi energy and the mobility edge can coincide; this is the disorder-induced Anderson transition. In many instances, one could justifiably say *all* instances, the individual effects of disorder must always be considered alongside electron–electron interactions (e.g. as in the Mott–Hubbard metal–insulator transition in, for example, the crystalline Hubbard model (Hubbard 1963, 1964*a, b*)).

The effects of such electron interactions are known to be particularly strong in narrow-band systems and cause them to be insulating when they should be metallic according to conventional band theory (Mott 1949, 1961). Well-known examples are NiO and LaCuO₄. The same basic problem can also be identified with all experimental systems traversing the metal–insulator transition. Typical examples would be the highly expanded states of alkali metals, and doped semiconductors (for reviews see Edwards & Rao 1985, 1995; Edwards *et al.* 1995).

To consider these issues within the context of these last two experimental systems, imagine, for example, a highly expanded alkali metal, such as rubidium, or a hydrogenic (donor) impurity atom of phosphorus substitutionally doped into a host semiconductor lattice (e.g. silicon). We discuss first the situation in which the lattice constant of the resulting assembly is so large that each rubidium atom or phosphorus donor behaves as an isolated entity which does not interact with other atoms or donors, respectively. To achieve genuine charge (electron) transport within this

assembly of neutral non-interacting atoms or donors requires the ionization of an outer valence electron from one of the sites and its transfer to another (neutral) site in the assembly, i.e. the formation of charged ionic states. This aspect is formally embodied within the so-called Mott–Hubbard correlation energy (U), which is the magnitude of the energy difference between the ionization energy (I) and the electron affinity (EA) of an isolated atom or donor; this energy, U (equal to $I - EA$), then represents the extra energy cost of putting two electrons (instantaneously) on one atomic site (Hubbard 1963, 1964*a, b*; Mott 1974, 1990). Since the magnitude of U is substantial (and positive) in the limit of isolated non-interacting atoms or donors, compared to other characteristic energies (including temperature), this highly expanded state of the system is unquestionably an insulator, for which $\sigma(T = 0 \text{ K}) = 0$.

At the opposite physical extreme of a small lattice constant, in our two cited examples, this would correspond to elemental rubidium at high densities (figures 4 and 5) or heavily doped Si:P, there exists considerable orbital overlap between the individual atomic or donor centres and the intersite electron transfer (or hopping) is greatly facilitated. Such enhanced intersite transfer at small atom (donor)–atom (donor) separations is manifest in a broadening of the one-electron band widths (Δ). Eventually this developing electronic band may become so wide that the band width compensates for the Hubbard repulsion energy; under these conditions spontaneous ionization from a neutral atom into states at the bottom of the band then occurs and the material becomes a metal ($\sigma(T = 0 \text{ K}) \neq 0$). Within this model, the criterion for the metal–insulator transition is generally taken as $\Delta \simeq U$, with the precise form of the equality depending upon the details of the geometry of the lattice of one-electron states (Hubbard 1963, 1964*a, b*; Berggren 1978; Mott 1974).

The Mott–Hubbard correlation energy has previously been used to attempt a ready demarcation of the naturally occurring elements of the periodic classification into metals and insulators (figure 8). For example, Friedel (1984) has used the values of U obtained from the difference between the first ionization energy of an atom and its first electron affinity. On this basis, metallic elements within the periodic system have values of the Hubbard U of about 10 eV or below, whereas atomic states for which U is in excess of *ca.* 8–9 eV tend to give rise to non-metallic (insulating) elements in the condensed phase under normal conditions. This is once again a good indication of the importance of atomic properties in dictating the critical conditions for the metallic versus non-metallic status of elements of the periodic classification.

Similar considerations can also be applied to chemical compounds, for example, binary and ternary transition metal oxides (Morin 1958; Goodenough 1963, 1971). The qualitative variation of the electron band width within the chemically similar transition-metal monoxides (MO, M = Ti through Mn) was recognized some time ago (Morin 1958) prior to any detailed quantum-mechanical calculations of the problem. Within the isostructural transition metal monoxide series one sees an emerging metal–insulator transition arising from the competition between Δ and U , but this time with the transition initiated by *chemical* variations in these physical parameters, with lattice constants showing relatively little change from TiO ($d_{\text{Ti-Ti}} \sim 2.94 \text{ \AA}$) to MnO ($d_{\text{Mn-Mn}} \sim 3.14 \text{ \AA}$). Stoichiometric TiO appears to exhibit the properties of itinerant electrons while the later members of the monoxide series have physical properties characteristic of localized electrons. Whether the d-electrons in these and other transition metal oxides are localized or itinerant now depends critically upon the magnitude of Δ and U for d-orbitals on neighbouring cations (Goodenough 1971).

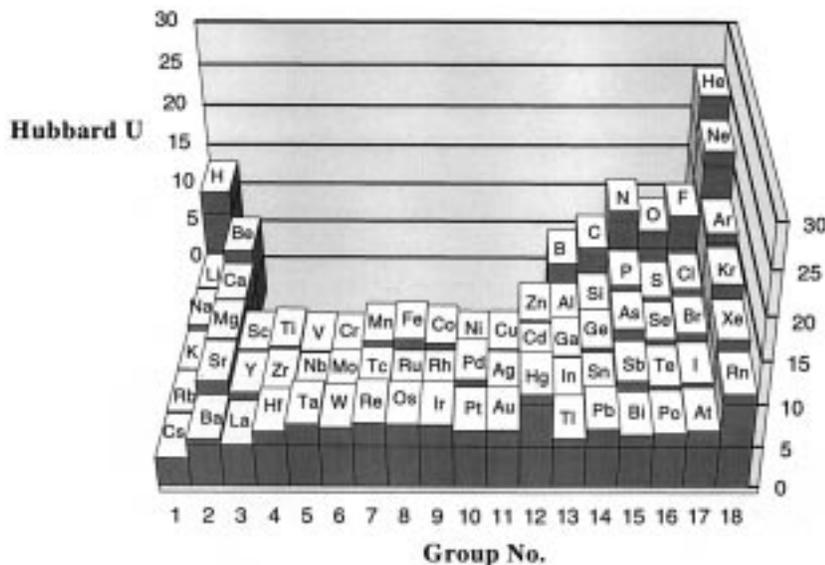


Figure 8. A three-dimensional plot of the Hubbard U ($= IE - EA$), in eV, for various atoms of the periodic classification of the elements. It can be seen that large (high) values of U correspond to insulating elements at standard conditions whilst low values of U (typically below 8–10 eV) delineate metallic elements.

In the first-row transition-metal monoxides, the slightly increased metal–metal internuclear distance (e.g. from TiO to MnO), and the anticipated 3d orbital contraction is expected to reduce the electronic band width because the overlap of the constituent wave functions will be significantly diminished. The combination of these effects will result in an increase in the effective mass of the 3d electrons and a lowering of their mobility. When the 3d band becomes extremely narrow (figure 9) it is no longer meaningful to assign a width to the band, and the charge carriers can be best considered to occupy discrete energy levels localized on the transition metal cations (Morin 1958). It was first pointed out by de Boer & Verwey (1937) and Mott (1937, 1949) that this situation must exist in oxides such as MnO, CoO and NiO, which are insulators when pure and stoichiometric and have room temperature resistivities exceeding $10^{10} \Omega \text{ cm}$. Morin (1958) showed that a 3d band having appreciable width exists in TiO and VO, whereas in the remaining transition-metal monoxides the corresponding 3d wavefunctions are localized.

Thus, the so-called ‘NiO problem’, a paradox for over half a century (Mott 1937; de Boer & Verwey 1937), can now be reconciled from a consideration of such periodic trends. For example, the well-established 3d orbital contraction across the transition metal series leads to a drastic reduction in the electron band width and a concomitant increase in the (intrasite) Hubbard correlation energy (Morin 1958; Goodenough 1971).

As Mott has repeatedly stressed, the simultaneous occurrence of, and interplay between, disorder and electron interactions is of cardinal importance; the true nature of the metal–insulator transition cannot be understood in terms of either effect in isolation (Mott 1974, 1985, 1990). This interplay between competing effects lies at the very heart of the phenomenon. For example, in a disordered strongly interacting system, Anderson localization due to disorder may tend to increase any local electron interaction effect (Ramakrishnan 1995). The problem of disordered interacting

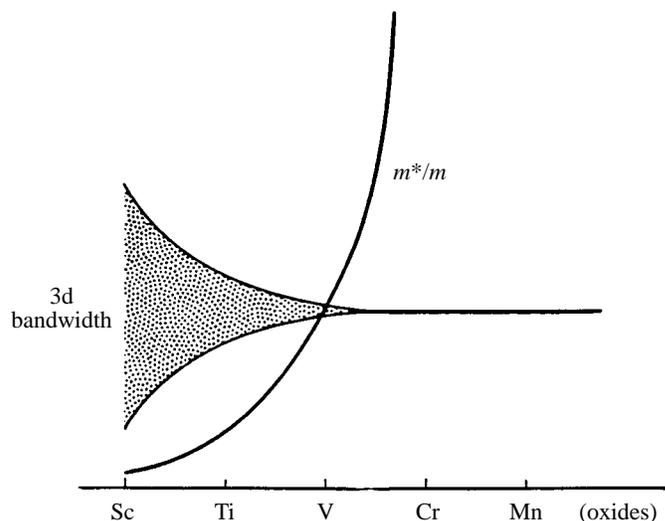


Figure 9. A schematic representation of the variation in the 3d electronic bandwidth, and the associated effective mass, m^*/m , of charge carriers for the transition element monoxides ScO through MnO. The underlying variations in bandwidth and m^*/m relate to the orbital contraction of the valence 3d wavefunctions with increasing atomic number; this has a dramatic effect on both Δ and U as we have a transition from the metallic (ScO) to the insulating (MnO), magnetic regime. Taken from Morin (1958).

systems at finite temperatures, is thus central to the nature of the metal–insulator transition (Logan *et al.* 1995). The combination of figure 2, and the sketch shown in figure 10, outlines the scale of the problem as applied to a real experimental system. We see in the latter a two-dimensional representation of the situation in a high-dielectric constant doped semiconductor (e.g. Si:P) at a donor concentration just below n_c (Holcomb 1995). The large physical dimensions of the isolated donor or impurity ‘atom’ (approximately 17 Å in Si:P) in comparison with the nearest-neighbour separation of two host silicon atoms (approximately 2.4 Å) means that, in relation to electronic structure, the system is far removed from any idealized description of a regular ordered lattice of hydrogen-like atoms, even though the material itself is perfectly crystalline. Indeed, the conducting metallic state is formed through an overlapping (random) percolating network of phosphorus donor atoms at a critical density of some 3.8×10^{18} atoms cm^{-3} . Since there clearly exist differing configurations of neighbours, with an inevitable spread of energy levels for electrons bound at various impurity sites, we have a considerable degree of disorder, in the Anderson sense. Similarly, electron–electron interactions (sketched out earlier) will always be present in this prototype highly disordered system.

This is the complex and fascinating reality of the experimental framework in which we must consider any Mott–Hubbard–Anderson models (Logan *et al.* 1955). It comes as no surprise, therefore, to find that the detailed understanding and theoretical description of this basic electronic phase transition remains a major challenge and a focus of attention worldwide.

(c) *Electrical conductivity at the metal–insulator transition; the minimum metallic conductivity*

In any discussion relating to the nature of the metal–insulator transition, it is important to say something about the anticipated value of the electrical conductivity

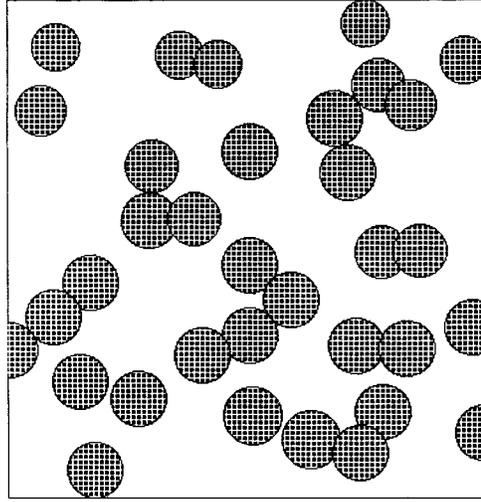


Figure 10. A schematic two-dimensional representation of a high-dielectric constant, doped semiconductor (Si:P) at an impurity concentration just below n_c , the critical density for the metal-insulator transition. The radius of each shaded circle is approximated as the Bohr radius (a_H^*) of the impurity, or donor phosphorus atom. The impurity system clearly forms a randomly overlapped state at this composition close to the critical transition region. Taken from Holcomb (1995).

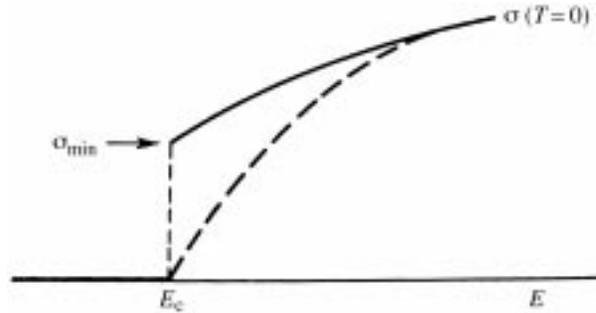


Figure 11. A schematic illustration of the two possibilities of a continuous versus discontinuous metal-insulator transition at zero temperature. The electrical conductivity should vanish at the point at which the Fermi level of the system passes through the localization threshold E_c . But does it vanish discontinuously (the solid curve) or smoothly (the dashed curve)? Here, the minimum metallic conductivity, σ_{\min} , at the transition is also shown.

at the transition point. As with the precise location of the transition, this matter continues to be the subject of intense debate.

Mott (1961) first proposed that the metal-insulator transition in a perfect crystalline material at $T = 0$ K is discontinuous (figure 11). He further argued that, *at the transition*, there exists a minimum conductivity, σ_{\min} (Mott 1972), for which the system can still justifiably be viewed as metallic, prior to the complete localization of the gas of itinerant conduction electrons (figure 1).

Mott's proposal (Mott 1972, 1982) was based on important arguments developed earlier by Ioffe & Regel (1960) in relation to the breakdown of the theory of electronic conduction in disordered semiconductors. Thus, in this Mott-Ioffe-Regel viewpoint, conventional Boltzmann transport theory becomes meaningless when the characteristic mean free path, l , of the itinerant conduction electrons becomes comparable

to, or less than, the interatomic spacing, d . In fact, this assertion derived from the proposal that the mean free path, according to its actual physical meaning, cannot be shorter than the electron wavelength (approximately k_F^{-1}), where k_F is the wave vector at the Fermi surface.

The Mott–Ioffe–Regel mean free path, l_{MIR} , under conditions of the minimum metallic conductivity, is equal to d . At this critical condition we have $d = d_c$. It then follows from Mott (1974), that the electrical conductivity of the metallic state at the metal–insulator transition (figure 11) cannot be smaller than the quantity σ_{min} , where

$$\sigma_{\text{min}} = C_{\text{Mott}}(e^2/h)d_c^{-1}, \quad (2.4)$$

here C_{Mott} is a constant which includes specific considerations relating to disorder, etc., and has a value in the region 0.01–0.05.

Mott’s concept, in essence, is that at zero temperature, the electrical conductivity of the metallic state continuously decreases with increase of disorder, and, upon reaching the value given by σ_{min} , the conductivity drops discontinuously to zero. Thus at zero temperature an itinerant (‘metallic’) conduction electron gas cannot possess a value of the conductivity less than the appropriate minimum possible value.

Note, however, that σ_{min} will be system different. This arises because d_c (approximately $n_c^{-1/3}$), from the Mott criterion (equation (2.1)) is unique for each system. When the critical conditions for metallization are reached, all carriers become itinerant. Clearly a system such as Si:P with $n_c \sim 3 \times 10^{18}$ electrons cm^{-3} will exhibit a lower minimum metallic conductivity than, say, expanded rubidium, for which $n_c \sim 10^{21}$ – 10^{22} electrons cm^{-3} . This important issue is further amplified shortly.

Abrahams *et al.* (1979) have, however, predicted a continuous metal–insulator transition on the basis of a scaling theory of non-interacting electrons in a disordered system, and their results question the very existence of σ_{min} in both two and three dimensions. The two possible scenarios for the form of the metal–insulator transition at zero temperature are compared in figure 11. The electrical conductivity at $T = 0$ K should vanish at the point at which the Fermi energy passes through the localization threshold, E_c at the mobility edge. The fundamental question, still unresolved after almost half a century of intense study, is whether the electrical conductivity vanishes discontinuously (solid curve) or smoothly (the dashed curve) at the transition? Today, a significant proportion of researchers appear to trust in the idea of a continuous metal–insulator transition (Edwards & Rao 1995); this, of course, within a model of non-interacting electrons. The concept of a discontinuous transition, and the necessary existence of a minimum metallic conductivity, may still be appropriate when electron–electron interactions are taken into account. Given the ongoing controversy, Thouless’s unfortunate comment (1982) that σ_{min} has been ‘one of the creative errors that have helped the progress of science’ now seems to be, at best, ill-judged and premature.

In spite of these difficulties, the concept of a minimum metallic conductivity continues to serve as a particularly useful experimental criterion for the metal–insulator transition in what may be called ‘the high temperature limit’ (Edwards *et al.* 1995; Rao 1996). The electrical conductivity at the metal–insulator transition, and n_c the critical density of carriers for such a transition can be related via

$$\sigma_{\text{min}} = C_{\text{Mott}}(e^2/h)n_c^{-3}. \quad (2.5)$$

In figure 12 we show the variation of σ_{min} with n_c for a range of systems, including the high temperature superconducting layered cuprates (Edwards *et al.* 1995; Rao

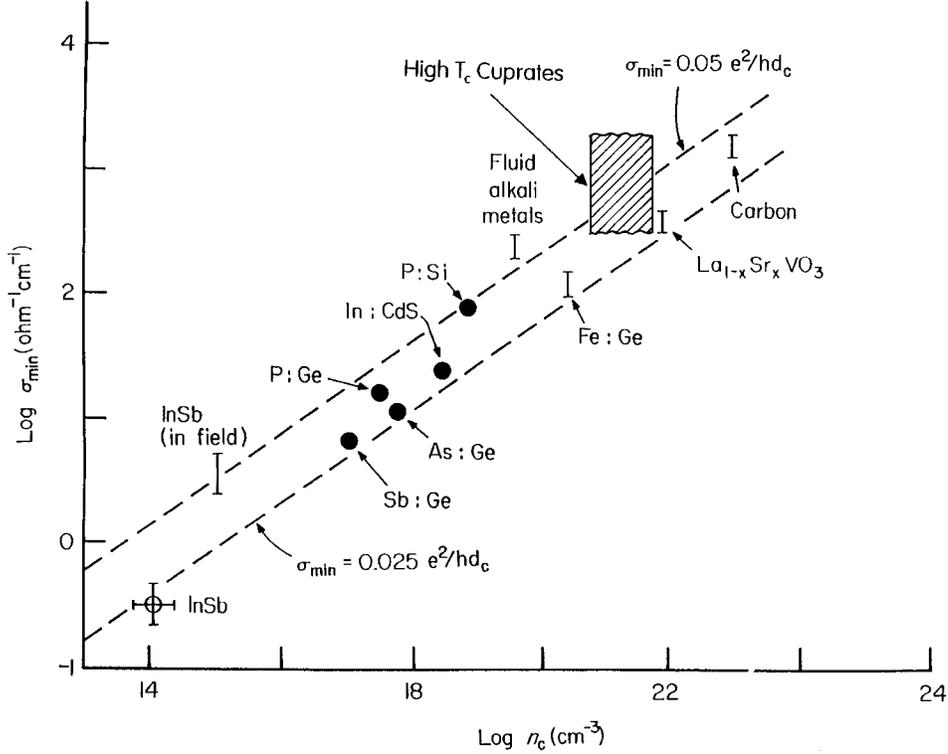


Figure 12. A (log-log) plot of the minimum metallic conductivity, σ_{\min} , against the critical carrier density, n_c , at the metal-insulator transition. The two straight lines correspond to values of $\sigma_{\min} = 0.05(e^2/h)d_c^{-1}$, and $\sigma_{\min} = 0.025(e^2/h)d_c^{-1}$. Adapted from Fritzsche (1978) and Mott (1982), in Edwards & Sienko (1983). We also include recent data for high-temperature superconducting cuprates (Rao 1996; Edwards *et al.* 1998).

1996). As first pointed out by Fritzsche (1978), σ_{\min} appears to represent satisfactorily the value of electrical conductivity in a wide range of materials for experimental conditions at which the activation energy for electrical conduction disappears.

3. Concluding remarks

The metal-insulator transition, a quantum phase transition at $T = 0$ K, is caused and accompanied by a fundamental qualitative change in electronic structure (figure 1). From common experience, this is manifestly self-evident. However, our detailed understanding of this most basic electronic transition, the transformation of metal to insulator, is still far from complete. Over 25 years ago Austin & Mott (1970) noted ‘... there is as yet no generally recognized theory of what happens at the transition point’. Even today this statement is rigorously correct. The metal-insulator transition, a subject initiated, inspired and led for over half a century by Sir Nevill Mott, continues to be one of the foremost intellectual challenges of condensed matter science.

References

Abrahams, E., Anderson, P. W., Liccardello, D. C. & Ramakrishnan, T. V. 1979 Scaling theory of localization: absence of quantum diffusion in two dimensions. *Phys. Rev. Lett.* **42**, 673–676.

Phil. Trans. R. Soc. Lond. A (1998)

- Anderson, P. W. 1958 Absence of diffusion in certain random lattices. *Phys. Rev.* **109**, 1492–1505.
- Ashcroft, N. W. 1993 The metal–insulator transition: complexity in a simple system. *J. Non-Crystalline Solids* **156**, 621–630.
- Austin, I. G. & Mott, N. F. 1970 Metallic and non-metallic behaviour in transition metal oxides. *Science* **168**, 71–77.
- Bardeen, J. 1940 Electrical conductivity of metals. *J. Appl. Phys.* **11**, 88–111.
- Berggren, K.-F. 1974 On the Herzfeld theory of metallization. *J. Chem. Phys.* **60**, 3399–3402.
- Berggren, K.-F. 1978 In *The metal–non-metal transition in disordered systems* (19th Scottish Universities Summer School in Physics, Edinburgh) (ed. L. R. Friedman & D. P. Tunstall), pp. 399–423.
- Cottrell, A. H. 1992 Will anybody ever use metals and alloys again? In *The science of new materials* (ed. A. Briggs), pp. 4–31. Oxford: Blackwell.
- Cottrell, A. H. 1997 In *Int. Centennial Symp. on the Electron* (ed. C. J. Humphreys & K. Takayanagi). (In the press.)
- de Boer, J. H. & Verwey, E. J. W. 1937 Semiconductors with partially and with completely filled 3d-lattice bands. *Proc. Phys. Soc. Lond.* **49**, 59–71.
- Edwards, P. P. 1995 The metal–non-metal transition: the continued/unreasonable effectiveness of simple criteria at the transition. In *Perspectives in solid state chemistry* (ed. K. J. Rao), pp. 250–271. New Delhi: Narosa.
- Edwards, P. P. & Rao, C. N. R. (eds) 1985 *The metallic and non-metallic states of matter*. London: Taylor and Francis.
- Edwards, P. P. & Rao, C. N. R. (eds) 1995 *Metal–insulator transitions revisited*. London: Taylor and Francis.
- Edwards, P. P. & Sienko, M. J. 1978 Universality aspects of the metal–nonmetal transition in condensed phases. *Phys. Rev. B* **17**, 2575–2581.
- Edwards, P. P. & Sienko, M. J. 1982 The transition to the metallic state. *Acc. Chem. Res.* **15**, 87–93.
- Edwards, P. P. & Sienko, M. J. 1983 What is a metal? *Int. Rev. Phys. Chem.* **3**, 83–137.
- Edwards, P. P., Ramakrishnan, T. V. & Rao, C. N. R. 1995 The metal–non-metal transition: a global perspective. *J. Phys. Chem.* **99**, 5228–5239.
- Edwards, P. P., Mott, N. F. & Alexandrov, A. S. 1998 The insulator–superconductor transformation in cuprates. *J. Supercond.* **11**, 151–154.
- Eggen, B. R., Murrell, J. N. & Dunne, L. J. 1997 Hydrogen, the first alkali metal? *Sol. St. Commun.* (In the press.)
- Ehrenreich, H. 1967 The electrical properties of materials. *Sci. Am.* **217**, 195–204.
- Feynman, R. P., Leighton, R. B. & Sands, M. 1964 *The Feynman lectures on physics*, II–1–2.
- Freyland, W. & Hensel, F. 1985 The metal–non-metal transition in expanded metals. In *The metallic and non-metallic states of matter* (ed. P. P. Edwards & C. N. R. Rao), pp. 93–121. London: Taylor and Francis.
- Friedel, J. 1984 In *Physics and chemistry of electrons and ions in condensed matter* (ed. J. V. Acrivos, N. F. Mott & A. D. Yoffe), p. 45. Dordrecht: Reidel.
- Fritzche, H. 1978 In *The metal–nonmetal transition in disordered systems* (Scottish Universities Summer School in Physics, Edinburgh) (ed. L. R. Friedman & D. P. Tunstall), pp. 193–238.
- Goldhammer, D. A. 1913 *Dispersion und absorption des lichts*. Leipzig: B. G. Teubner.
- Goodenough, J. B. 1963 *Magnetism and the chemical bond*. New York: Interscience.
- Goodenough, J. B. 1971 Metallic oxides. *Prog. Solid State Chem.* **5**, 145–399.
- Hensel, F. & Edwards, P. P. 1996a Hydrogen: the first metallic element. *Science*.
- Hensel, F. & Edwards, P. P. 1996b Hydrogen, the first alkali metal. *Chem. Eur. J.* **2**, 1201–1203.
- Hensel, F. & Edwards, P. P. 1996c The changing phase of expanded metals. *Phys. World.* 43–46.
- Herzfeld, K. F. 1927 On atomic properties which make an element a metal. *Phys. Rev.* **29**, 701–705.

- Holcomb, D. F. 1995 The metal–insulator transition in doped semiconductors. In *Metal–insulator transitions revisited* (ed. P. P. Edwards & C. N. R. Rao), pp. 65–77. London: Taylor and Francis.
- Hubbard, J. 1963 Electron correlations in narrow energy bands. *Proc. R. Soc. Lond. A* **276**, 238–257.
- Hubbard, J. 1964a Electron correlations in narrow energy bands. II. The degenerate band case. *Proc. R. Soc. Lond. A* **277**, 237–259.
- Hubbard, J. 1964b Electron correlations in narrow energy bands. III. An improved solution. *Proc. R. Soc. Lond. A* **281**, 401–419.
- Ioffe, A. F. & Regel, A. R. 1960 Non-crystalline, amorphous and liquid electronic semiconductors. *Proc. Semicond.* **4**, 239–291.
- Iye, H. 1995 Superconductor–insulator transition in cuprates. In *Metal–insulator transitions revisited* (ed. P. P. Edwards & C. N. R. Rao), pp. 211–230. London: Taylor and Francis.
- Logan, D. E. & Edwards, P. P. 1985 The periodic system of the elements. In *The metallic and nonmetallic states of matter* (ed. P. P. Edwards & C. N. R. Rao), pp. 65–92. London: Taylor and Francis.
- Logan, D. E., Szezech, Y. H. & Tusch, M. A. 1995 Mott–Hubbard–Anderson models. In *Metal–insulator transitions revisited* (ed. P. P. Edwards & C. N. R. Rao), pp. 343–394. London: Taylor and Francis.
- McMillan, E. M. 1963 Private communication, cited in C. Kittel, *Introduction to solid state physics*, 2nd edn, p. 270. London: Wiley.
- Morin, F. J. 1958 *Oxides of the 3d transition metals*. In *Semiconductors* (ed. N. B. Hannay), pp. 600–633. New York: Reinhold.
- Mott, N. F. 1937 Discussion of the paper by J. H. de Boer and E. J. W. Verwey. *Proc. Phys. Soc. Lond.* **49**, 72–73.
- Mott, N. F. 1949 The basis of the electron theory of metals, with special reference to the transition metals. *Proc. Phys. Soc. Lond. A* **62**, 416–422.
- Mott, N. F. 1956 On the transition to metallic conduction in semiconductors. *Can. J. Phys.* **34**, 1356–1368.
- Mott, N. F. 1958 The transition from the metallic to the non-metallic state. *Nuovo Cimento*, vol. VII, series X, no. 2, pp. 312–328.
- Mott, N. F. 1961 The transition to the metallic state. *Phil. Mag.* **6**, 287–309.
- Mott, N. F. 1972 Conduction in non-crystalline systems. IX. The minimum metallic conductivity. *Phil. Mag.* **26**, 1015–1026.
- Mott, N. F. 1974 *Metal–insulator transitions*. London: Taylor and Francis.
- Mott, N. F. 1982 Metal–insulator transitions. *Proc. R. Soc. Lond. A* **382**, 1–24.
- Mott, N. F. 1985 Metals, non-metals and metal–non-metal transitions. In *The metallic and nonmetallic states of matter* (ed. P. P. Edwards & C. N. R. Rao), pp. 1–21. London: Taylor and Francis.
- Mott, N. F. 1990 *Metal–insulator transitions*. London: Taylor and Francis.
- Pauling, L. 1938 The nature of the interatomic forces in metals. *Phys. Rev.* **54**, 899–904.
- Pauling, L. 1984 The metallic orbital and the nature of metals. *J. Solid State Chem.* **54**, 297–307.
- Ramakrishnan, T. V. 1995 Metal–insulator transitions—whither theory? In *Metal–insulator transitions revisited* (ed. P. P. Edwards & C. N. R. Rao), pp. 293–307. London: Taylor and Francis.
- Rao, C. N. R. 1996 Virtues of marginally metallic oxides. *Chem. Commun.*, 2217–2222.
- Thouless, D. J. 1982 In *Anderson localization* (ed. Y. Nagayoka & Fukuyama), p. 2. Berlin: Springer.
- Weir, S. T., Mitchell, A. C. & Nellis, W. J. 1996 Metallization of fluid molecular hydrogen at 140 GPa (1.4 Mbar). *Phys. Rev. Lett.* **76**, 1860–1863.
- Wigner, E. & Huntington, H. B. 1936 On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.* **3**, 764–770.