PERSPECTIVE

Dalton
www.rsc.org/dalton

Basic materials physics of transparent conducting oxides †

P. P. Edwards,[a] A. Porch,[b] M. O. Jones,[a] D. V. Morgan[b] and R. M. Perks[b]

[a] *Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford, UK OX1 3QR*

[b] *School of Engineering, Cardiff University, Queen's Buildings, The Parade, PO Box 925, Cardiff, UK CF24 0YF*

Received 10th June 2004, Accepted 29th June 2004
First published as an Advance Article on the web 6th August 2004

Materials displaying the remarkable combination of high electrical conductivity and optical transparency already from the basis of many important technological applications, including flat panel displays, solar energy capture and other opto-electronic devices. Here we present the basic materials physics of these important materials centred on the nature of the doping process to generate n-type conductivity in transparent conducting oxides, the associated transition to the metallic (conducting) state and the detailed properties of the degenerate itinerant electron gas. The aim is to fully understand the origins of the basic performance limits of known materials and to set the scene for new or improved materials which will breach those limits for new-generation transparent conducting materials, either oxides, or beyond oxides.

† Based on the presentation given at Dalton Discussion No. 7, 5–7th July 2004, University of St Andrews, UK.

1 Preamble

The electron-gas model of solids applies equally to excess intinerant electrons in the conduction band of both metals and semiconductors. Within the Boltzmann formulation, the electrical conductivity of such a material is given by $\sigma = ne\mu$ where n is the density of free carriers, μ is their mobility and e is the electronic charge. The mobility is directly proportional to the free carrier resistivity relaxation time, τ (the time between resistive scattering events), and is inversely proportional to the carrier effective mass, m^*, *via* $\mu = e\tau/m^*$. It is possible to establish a materials 'sorting map' of electrical conductors within an 'n–μ' diagram, and such a representation is given in Fig. 1. Here constant conductivity contours are the straight lines shown. It is immediately apparent that a number of 'clusters' are readily discernable within this materials-space; namely semimetals (top-left), the prototypical, highly conducting metals (upper-middle) and conventional semi-conductors (bottom-right). There is another key class of materials, the so-called transparent conducting oxides, the basis of this presentation, which

Peter P. Edwards

Peter P. Edwards FRS is Professor and Head of Inorganic Chemistry at Oxford. A graduate of Salford University (BSc and PhD) he spent several periods at Cornell, Cambridge and Birmingham before assuming his current position in October 2003. His research interests span the metal–insulator transition; Edwards is also coordinator of the UK Sustainable Hydrogen Energy Consortium (UK-SHEC; www.uk-shec.org).

Adrian Porch

Adrian Porch is a Reader in the School of Engineering at Cardiff University. He received a BA (Hons) degree in Physics from the University of Cambridge in 1986 and PhD from the University of Cambridge in 1991 for research into the fundamental properties of high-temperature superconductors at microwave frequencies. From 1990 to 2000 he was an academic in the Electronic Materials and Devices Group in the School of Electronic and Electrical Engineering at the University of Birmingham. He joined Cardiff University in the Autumn of 2000. Dr Porch's research interests concern the fundamental properties and applications at high frequencies of novel electronic materials over a wide range of temperatures and frequencies; these materials include superconductors, low-dimensional conductors, composite radar absorbers and low loss dielectrics. He has authored and co-authored over 60 international journal papers and over 50 conference papers in these areas.

DOI: 10.1039/b408864f

Martin Owen Jones

Martin Owen Jones is a Research Assistant in the Inorganic Chemistry Laboratory at Oxford University. He received a BSc (Hons) degree in chemistry from Nottingham University in 1989 and a PhD in 1993 for research into complex ternary oxides. His research interests concern the synthesis, properties and applications of bulk and thin film functional materials.

David Vernon Morgan

David Vernon Morgan (F.R.Eng) is currently a Distinguished Research Professor (Microelectronics) at Cardiff University. He has previously been the Head of the School of Engineering (1992–2002). He has published over 220 scientific papers, edited 10 research monographs and co-written two text books on electronic devices. He was elected to the fellowship of the Royal Academy of Engineering in 1996.

Richard Marc Perks

Richard Marc Perks is a Lecturer at Cardiff University School of Engineering. His research interests include semiconductor device processing, periodic structures and bio-metrics. He has previously been a senior research scientist at British Aerospace and is a member of the Institute of Physics.

exhibit intermediate characteristics between the highly conducting/reflecting elemental metals and the low carrier density/high-mobility prototypical semiconducting materials such as Si or GaAs. Importantly this class of materials, typically indium tin oxide, tin oxide and indium oxide are characterised by relatively high values of both the electron density and the electron mobility sufficient for them to exhibit substantial electrical conductivity as transparent solids, most notably in thin film form.

2 Transparent conducting oxides

Sn-doped In_2O_3 thin films [$In_{2-x}Sn_xO_3$:ITO] exhibit a remarkable combination of optical and electrical transport properties: *viz*. (i) high electrical conductivity (~$10^4\ \Omega^{-1}\ cm^{-1}$) and (ii) high optical transparency (>80%) in the visible range of the spectrum.[1–9] These seemingly contradictory properties of close-to 'metallic' conductivity in a material simultaneously exhibiting almost complete 'non-metallic' or insulating transparency form the basis of numerous important applications in contemporary and emerging technologies. A typical example is the use of transparent current-spreading layers in light-emitting diodes where a low sheet resistance can significantly enhance the optical power output of the device.[10,11] Similarly if the electrical conductivity of a transparent conductor is sufficiently high, the thermal emittance is accordingly low, which opens up many applications for the thermal insulation of windows and other manufactured structures.[1–3]

The literature on ITO and related transparent conductors is vast, testifying to more than 50 years of intensive scientific investigation and technical application. However, as pointed out by Tanaka *et al*.[12] the guiding principles for the materials/processing design of these materials still represents a challenging goal. In certain respects the basic physics of this, arguably the rarest form of conductivity – electronic conduction in a transparent solid – is now reasonably well understood. Thus, the high conductivity of ITO is related to the presence of shallow donor or impurity states located close to the host (In_2O_3)

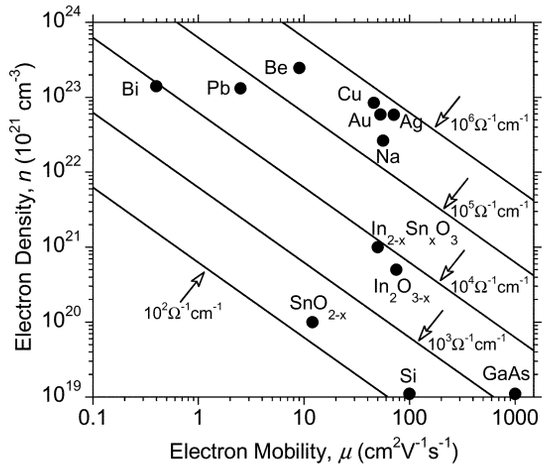

Fig. 1 *Sorting* (*Conducting*) *solids*: materials space for semimetals, good metals, transparent conductors and semiconductors based on n–μ correlations (electron carrier density and electron mobility, respectively). Data shown are for room-temperature measurements. Constant conductivity contours are shown as straight lines.

conduction band, and produced *via* chemical doping of Sn^{4+} for In^{3+} or by the presence of oxygen vacancy impurity states in In_2O_{3-x}. The proximity of any such electronic states to the host conduction band necessitates their classification as shallow donor or impurity states, whereby the excess or donor electron undergoes thermal ionization at room temperature into the host conduction band. Ultimately, upon further doping, a degenerate gas of current-carrying electrons is created which gives rise to far-infrared (Drude-like) absorption and high electronic conductivity, but at the same time the fundamental host band gap is left intact, *i.e.* the electrically conductive material remains optically transparent in the visible region. So, given that this is a reasonable description of the state-of-events, is the chapter on electronic transparent conductors now closed? No – far from it – the never-ending demands from technological applications for better quality electronic materials exhibiting higher performance limits continues to inspire research worldwide toward a deeper understanding of the 'materials physics' of the microscopic electronic structure and mechanisms driving the unique combination of electrical and optical properties in transparent conducting oxides. At the same time, advances in the 'materials chemistry' of transparent conductors – namely, advances in our fundamental understanding of the chemical and structural origins of such materials – continues apace.[5–7] This avenue also shows great promise for the development of synthetic chemical routes to dramatically improved materials or potentially new materials, possibly even for existing materials. There is another reason why this conjoint materials chemistry and physics approach is key; namely, it appears that within existing transparent conducting oxide materials the intrinsic performance limits are now regularly being approached. A fundamental understanding of how such an unusual combination of electronic and optical properties can arise, targeted in a way as to understand and impact upon the performance limits of these materials, must underpin any modern programme of research aimed at improved transparent conducting materials. Specifically, the elucidation of electronic structure, the nature of the (composition-induced) electronic phase transition to the degenerate itinerant electron gas (the nonmetal-to-metal transition) and the nature of the itinerant-electron scattering mechanisms once in the degenerate electron gas regime, all form part of the necessary ingredients for chemically engineering materials – oxides and beyond oxides – with potentially improved performance characteristics. Here we attempt to review just some of the basic concepts with a view to understanding the electronic structure and properties of these important materials. Our hope is that such an approach will be of value in combining theory and experimentation, essential ingredients for the optimization of materials properties.[6]

3 Elements of a theoretical framework for the metal–nonmetal transition

The following picture has now emerged for the conditions by which the degenerate gas of electrons is ultimately formed upon doping, for example, In_2O_3 by SnO_2. Semiconducting In_2O_3 is best regarded as a host oxide lattice to which SnO_2 can be added as a substitutional dopant; thus the metal atom of higher valency (Sn^{4+}) forms a singly charged donor state (nominally Sn^{3+}; Fig. 2) by substituting for an In^{3+} atom of the host material. In_2O_3 itself has a particularly complex crystal structure (*ca.* 80 atoms per unit cell) but the one-electron electronic energy band model originally developed by Fan and Goodenough[13] still represents an excellent and transparent starting description for describing the electronic structure of both the pure and doped materials.[14]

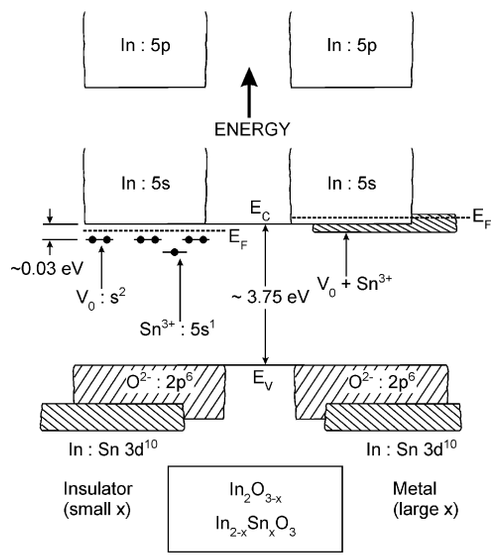

Fig. 2 Schematic energy-band model for tin doped In_2O_3 for small x (insulating) and large x (metallic) modified from Fan and Goodenough.[13] The issue as to whether the 'impurity band' (for large x) is separate from, or placed inside the In 5s (host) conduction band is not resolved at this time.

In_2O_3 has a filled valence band that is primarily oxygen 2p in character. The (filled) In $3d^{10}$ core levels also lie well below the valence band edge, as depicted in Fig. 2. The direct semiconductor bandgap, denoted E_g, is 3.75 eV. It has been proposed that the conduction band is mainly derived from the indium 5s electrons and for the undoped material the Fermi energy is located half way between the valence and conduction bands. With the substitution of In^{3+} by Sn^{4+}, one electron donor or impurity states are formed below the (host) In 5s or 5sp conduction band. First-principle electronic structure calculations on In_2O_3 reveal that the unoccupied In 5sp orbitals show wide spatial distribution, even extending over the third In shell. The electronic structure of this individual impurity state is best described in terms of Effective Mass Theory[15] in which the wave function of the donor state can be written:

$$\psi(r) = \sum_{j=1}^{N} \alpha_j F_j(r) \phi_j(r) \qquad (1)$$

where N is the number of equivalent minima in the E/k band structure, a_j are the relevant numerical coefficients, ϕ_j is the

Bloch (running wave) at the *j*th minima, and the functions $F_j(r)$ are 'hydrogen-like' envelope functions.

Remarkably, solutions of the appropriate Schrödinger wavefunction for this impurity state yield results very similar to the wave function for the isolated hydrogen atom, the difference being that the free electron mass m_0 of hydrogen is replaced by an effective mass m^* and the (ground state) Bohr radius, a_H^*, now replaces the hydrogen radius, a_0. The lowest eigenvalue has a mean Bohr radius of the order of

$$a_H^* = \frac{\hbar}{m^*\left(e^2/\varepsilon_r\right)} = \frac{a_0\varepsilon_r}{\left(m^*/m_0\right)} \qquad (2)$$

where m^* is an appropriate effective mass, a_0 is the normal (hydrogen) Bohr radius and ε_r relative permittivity.

Collected data from the ITO literature suggest that $m^* \simeq 0.35m_0$ and $\varepsilon_r \sim 4$, giving $a_H^* \simeq 13$ Å. This is a rather large orbit, extending over several hundred crystal cells, but, importantly, localised at the donor (substitutional) impurity atom.

Mott[16,17] first introduced the idea that a composition-induced transition from nonmetallic to metallic type conduction (and *vice-versa*) might occur abruptly at $T = 0$ K at a critical concentration of states, n_c which is given by

$$n_c^{1/3}a_H^* = 0.26 \pm 0.05 \qquad (3)$$

That a composition-induced transition can occur in the present materials can be readily discerned from the early variable-temperature resistivity data of Marley and Dockerty[18] for as-grown, heat-treated and antimony-doped crystals of SnO_2 shown in Fig. 3. We identify the metallic (itinerant electron) and non-metallic (localized electron) regimes by the electrical resistivity (or conductivity) data suitably extrapolated to $T \longrightarrow 0$, where one then attempts to classify materials as either metallic or nonmetallic. Nonmetallic (insulating) materials are defined as exhibiting zero conductivity at the absolute zero of temperature. In contrast, metallic materials always display finite conductivities at absolute zero,[16] *viz*;

$$\sigma(T \longrightarrow 0\ \text{K}) \neq 0\text{: Metallic} \qquad (4)$$

$$\sigma(T \longrightarrow 0\ \text{K}) = 0\text{: Nonmetallic} \qquad (5)$$

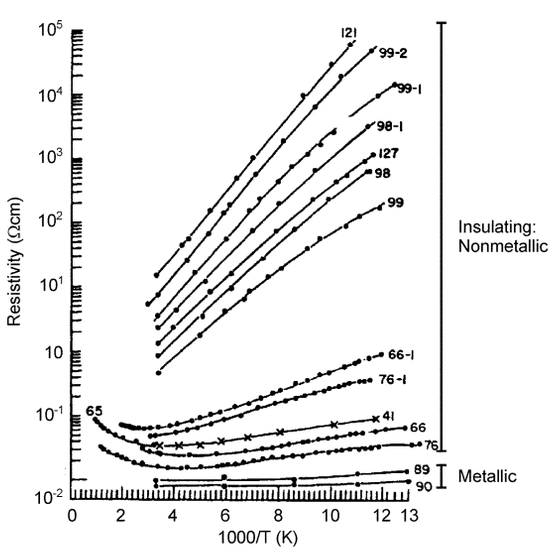

Fig. 3 Electrical resistivity as a function of $1/T$ between 80 and 900 K for as-grown, heat-treated and antimony-doped crystals.[18] The numbers at each curve represent sample labels of individual crystals. The temperature dependence of resistivity sets the metallic and insulating (nonmetallic) regimes.

Unfortunately the experimental data for the SnO_2:Sb_2O_5 system extends only down to temperatures of 80 K, but a composition-induced metal–nonmetal transition is clearly evident. The doping process can also be effected by oxygen vacancy creation, as in In_2O_{3-2x} corresponding to $InO_{3/2-x}\square_{1/2+x}$ where $\square$ symbolizes a vacant tetrahedral interstice.[13] Oxygen vacancies provide electrons by acting as doubly charged donors. Of these vacancies, one-half per In^{3+} ion are ordered and x per In^{3+} ion are randomly distributed over the O^{2-} ion subarray. These vacancies are generally symbolised by V_0, signifying oxygen-array vacancies (actually, as-prepared indium oxide is generally somewhat reduced). Fan and Goodenough[13] noted that each V_0 is surrounded by In^{3+} ion 5s orbitals that are stabilized from the In 5s conduction band by a lack of covalent bonding with the missing O^{2-}. Thus, In 5s orbitals at each V_0 site form shallow donor states just below E_c (the energy of the conduction band edge) that trap two electrons per oxygen vacancy and measurements suggest they lie *ca*. 0.03 eV below E_c for low x. At high V_0 concentrations (*i.e.* large x values, and low temperature) one anticipates either the overlap of the constituent electron donor state orbitals to form a discrete 'impurity band' or a situation whereby this V_0 impurity band forms and overlaps E_c at the bottom of the host conduction band. Either way, one anticipates an electronic transition to a degenerate, itinerant-electron semiconductor, *i.e.* a transition to what we define here as the metallic state (eqn. (4)). Here, there are once again many resonances to the situation in what one would term 'conventional' doped semiconductors (*e.g.* Si:P). Graham *et al.*[19] have investigated in detail the metal–nonmetal transition in amorphous In_2O_3 thin films in which they evaluate the metal–nonmetal transition in relation to the so-called Ioffe–Regel[20] criterion which predicts a metal–nonmetal transition when $k_F l_e \approx 1$ is satisfied. Here $k_F = (3\pi^2 n)^{1/3}$ is the Fermi wavenumber expressed in terms of the carrier density (n) and l_e is the elastic mean free path of the free carriers, with $l_e = v_F\tau = \hbar k_F\tau/m^*$, where v_F is the Fermi velocity and τ is the resisitivity relaxation time. Using the Boltzmann conductivity expression $\sigma = ne^2\tau/m^*$ to eliminate τ/m^*, one obtains for the parameter $k_F l_e = \hbar(3\pi^2)^{2/3}/(e^2\rho n^{1/3})$ where ρ is the resisitivity. The carrier density in these experiments was determined from voltage measurements (Fig. 4).

Mott further deduced that the value of the electrical conductivity when metallic conductivity first occurs, or equivalently the conductivity value when metallicity ceases, is;[16,17,21]

$$\sigma_{min} = 0.026e^2/\hbar d_c \qquad (6)$$

where d_c, the average distance between impurity centres at n_c, is about $2.5a_H^*$ according to eqn. (3). Note that the numerical constant in eqn. (6) may lie between 0.025 and 0.04 depending on the approximations used and on the average coordination number, z, of the centres.

The general applicability of the Mott criterion (eqn. (3)) has been demonstrated by Edwards and Sienko,[22] who assembled the data for doped semiconductors and other materials; we include in Fig. 5 the available experimental data for the metal–nonmetal transition in the transparent oxide conductors. We believe that this theoretical framework – set within that known for other materials and systems undergoing metal–nonmetal transitions – gives a reasonable working description of the composition-induced transition from insulating (nonmetallic) to metallic behaviour in the transparent conducting oxide materials. Similarly, analysis of resistivity–temperature curves (Figs. 3 and 4) gives a reasonable description of σ_{min} for these systems roughly in line with Mott's minimum metallic conductivity (Fig. 6 and ref. 21). However, more detailed investigations of the metal–nonmetal transition in both the established and the emerging new materials is urgently called for, particularly in view of the issues confronting the nature of the itinerant

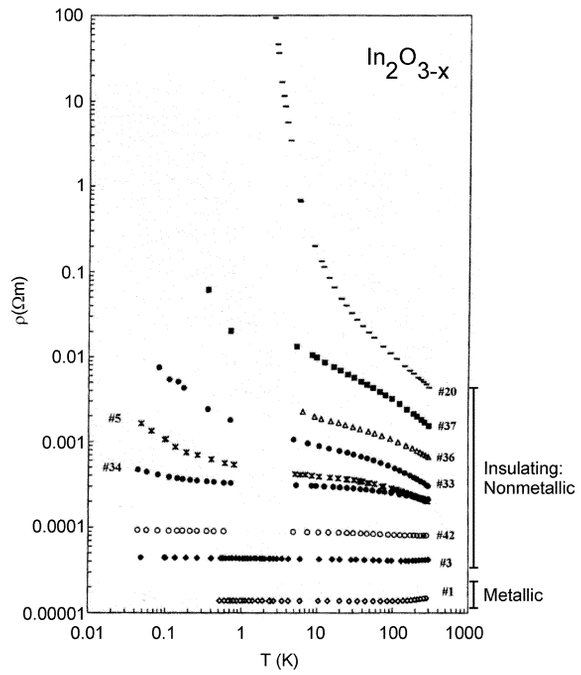

Fig. 4 Electrical resistivity *vs.* temperature dependence for a number of amorphous indium oxide samples. A clear metal–nonmetal transition is observed.[19] Sample #1 is metallic, all the others are shown to be nonmetallic by careful analysis of their temperature-dependent resistivity behaviour.

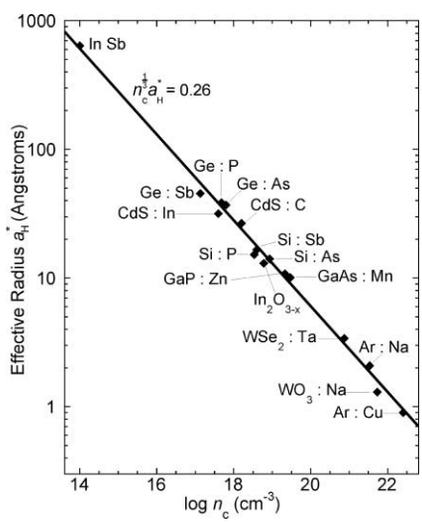

Fig. 5 Comparison of Mott's criteria for metal–non-metal transition with experimental data.[22] For In_2O_3 $a_H^* \simeq 13$ Å and $n_c \sim 6 \times 10^{18}$ cm^{-3}.

electron gas in the metallic regime (see below). Having sketched out the background of the composition-induced transition to the degenerate electron regime, we can now move to review the physics of that state.

4 Basic materials physics of the degenerate electron gas

The 'optical window' for ITO is set at short wavelengths by its direct band gap (around 3.75 eV) and at longer wavelengths by its plasma edge, which is in the near infrared provided that the free electron density, n, does not exceed 2×10^{21} cm^{-3} (as we will see below). For doping levels above that set by the Mott criterion[22] (*i.e.* $n > n_c \approx 10^{19}$ cm^{-3} for ITO; eqn. (3)), one can consider ITO to have a full valence band and a parabolic conduction band partially filled by a degenerate free electron gas[1,23] up to the Fermi energy, $E_F = \hbar^2(3\pi^2 n)^{2/3}/2m^*$ (here m^* is the effective mass of conduction band electrons).

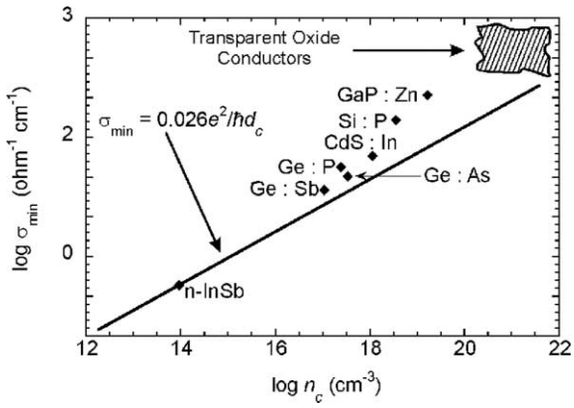

Fig. 6 Comparison of experimental data with Mott's minimum metallic conductivity σ_{min}. The data for transparent oxide conductors reveals the lack of precise estimates of σ_{min} in these systems at this point in time. The original data for doped semiconductors is from ref. 21.

Within this free electron framework the optical properties are described quite well by the simple Drude model but with a strongly frequency-dependent electron scattering time for Fermi surface electrons.[1] Electron scattering in ITO films occurs from a number of different scattering centres such as neutral impurities, phonons and defects, though the most dominant are likely to be the impurity ions responsible for the doping.[23]

A local current density $\boldsymbol{J}$ within a conducting material in response to a time-dependent electric field $\boldsymbol{E}$ is described well by the simple Drude model:

$$\frac{\mathrm{d}\boldsymbol{J}}{\mathrm{d}t} + \frac{1}{\tau}\boldsymbol{J} = \frac{ne^2}{m^*}\boldsymbol{E} \tag{7}$$

Defining the dynamic conductivity $\sigma(\omega)$ in the usual way *via* $J(\omega) = \sigma(\omega)E(\omega)$, we find that

$$\sigma(\omega) = \left(\frac{ne^2}{m^*}\right)\left(\frac{1}{1+i\omega\tau}\right) \equiv \frac{\sigma_0}{1+i\omega\tau} \tag{8}$$

where the dc conductivity is $\sigma_0 = ne^2\tau/m^* \equiv \omega_N^2\varepsilon_0\tau$, with $\omega^2 = ne^2/\varepsilon_0 m^*$. Since $\nabla \times \boldsymbol{H} = \boldsymbol{J} + i\omega\varepsilon_\infty\varepsilon_0\boldsymbol{E} \equiv i\omega\varepsilon(\omega)\varepsilon_0\boldsymbol{E}$, the Drude form of the frequency-dependent dielectric function $\varepsilon(\omega) = \varepsilon_1(\omega) - i\varepsilon_2(\omega)$ is

$$\varepsilon(\omega) = \varepsilon_\infty + \frac{\sigma(\omega)}{i\omega\varepsilon_0} = \varepsilon_\infty - \left(\frac{\omega_N^2\tau^2}{1+\omega^2\tau^2}\right) - \left(\frac{i\omega_N^2\tau}{\omega}\right)\left(\frac{1}{1+\omega^2\tau^2}\right) \tag{9}$$

where ε_∞ is the high-frequency relative permittivity of the ITO lattice. Separating the real and imaginary parts of $\varepsilon(\omega)$ yields

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_N^2\tau^2}{1+\omega^2\tau^2}, \quad \varepsilon_2(\omega) = \frac{1}{\omega}\left(\frac{\omega_N^2\tau}{1+\omega^2\tau^2}\right) \tag{10}$$

Collected ITO data from the literature[1] are in agreement that $\varepsilon_\infty \approx 4.0$ and $m^* \approx 0.35m_e$. Electron mobilities reported for high quality ITO films are around $\mu_e \approx 50$ cm^2 V^{-1} s^{-1},[9,23] giving a dc scattering time $\tau_0 \approx m^*\mu_e/e \approx 10^{-14}$ s. Drude fits to the optical data of these films suggest that $\tau \approx 3.3 \times 10^{-15}$ s,[1] somewhat lower than the value of τ_0 inferred from μ_e, so in our calculations we assume that $\tau = \tau_0/3$. The important assumption here is that $\mu_e \propto \tau_0 \propto \tau$; the scaling factor relating τ and τ_0 does not affect our general conclusions.

The condition $\varepsilon_1(\omega_p) = 0$ defines the plasma frequency ω_p of a conducting material and when $\omega < \omega_p$ a conductor is highly reflective, since then $\varepsilon_1 < 0$. The plasma frequency may be written:

$$\omega_p = \left(\frac{{\omega_N}^2}{\varepsilon_\infty} - \frac{1}{\tau^2} \right)^{1/2} \quad (11)$$

For ITO in the limit of high electron density (*i.e.* $n > 10^{21}$ cm^{-3}) and high mobility (*i.e.* $\mu_e = e\tau_0 m^* > 20$ cm^2 V^{-1} s^{-1}) then ${\omega^2}_N\tau^2/\varepsilon_\infty \gg 1$ and consequently:

$$\omega_p \approx \frac{\omega_N}{\sqrt{\varepsilon_\infty}} = \left(\frac{ne^2}{\varepsilon_\infty \varepsilon_0 m^*} \right)^{1/2} \quad (12)$$

Hence, in this limit $\omega_p \propto \sqrt{n}$ and thus is determined almost exclusively by n and not by τ (*i.e.* electron mobility μ_e, recalling that $\mu_e = e\tau_0/m^*$). This means that for a film to be non-reflective to light of free-space wavelength λ_0 (*i.e.* $\omega_p < 2\pi c/\lambda_0$), the electron density n must satisfy:

$$n < \frac{4\pi^2 \varepsilon_\infty m^*}{\mu_0 e^2 {\lambda_0}^2} \quad (13)$$

which for ITO yields the condition $n(\text{cm}^{-3}) < 1.6 \times 10^{27}/\lambda_0^2$ (nm^2).

Therefore, for efficient transmission of the whole visible spectrum (including red light of free-space wavelength up to 780 nm), the electron density should not exceed 2.6×10^{21} cm^{-3}. To allow a suitable 'safety margin', we'll consider a plasma wavelength just over 1 μm (*i.e.* in the near-infrared), in which case the maximum tolerable free electron density is about $n_{max} \approx 1.5 \times 10^{21}$ cm^{-3}. This value has already been exceeded in various ITO films, but clearly values of $n > n_{max}$ are detrimental to the transparency of the films owing to the plasma edge creeping into the red part of the visible spectrum. The power reflection coefficient R at an interface between a thick ITO layer and free space can now be calculated from

$$R(\omega) = \left| \frac{\sqrt{\varepsilon}(\omega) - 1}{\sqrt{\varepsilon}(\omega) + 1} \right|^2 \quad (14)$$

This takes the following limiting forms for frequencies below and well above the plasma frequency

$$R \approx 1 - 2\sqrt{\frac{2\omega}{{\omega_N}^2 \tau}} \approx 1 - 2\sqrt{\frac{2\omega\varepsilon_0}{\sigma_0}}, \ (\omega < \omega_p) \quad (15)$$

$$R \approx \left| \frac{\sqrt{\varepsilon_\infty} - 1}{\sqrt{\varepsilon_\infty} + 1} \right|^2 \approx 0.11, \ (\omega \gg \omega_p) \quad (16)$$

An exact calculation of R using the full Drude form for $\varepsilon(\omega)$ is shown in Fig. 7 as a function of n for various incident wavelengths in the visible spectrum. For frequencies below the plasma frequency then R is close to 1 (*i.e.* it is very reflective), with a small amount of electromagnetic absorption within the material (approximately equal to $1 - R$). We now calculate the electromagnetic skin depth, which determines ultimately the transparency of a conducting film since the normalised power transmission coefficient $T(t)/T(0)$ for a film of finite thickness t is very well approximated by the function $\exp(-2t/\delta)$. For example, the quantity $T(t)/T(0)$ for a film of thickness $t = \delta$ is

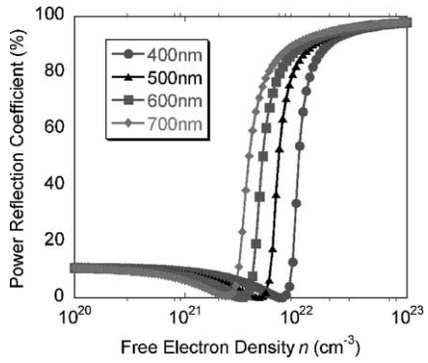

Fig. 7 The power reflection coefficient at the interface between a thick ITO layer and free space as a function of free electron density n for various incident wavelengths in the visible spectrum. We assume the limit ${\omega_N}^2\tau^2/\varepsilon_\infty \gg 1$, meaning that the plasma frequency is determined almost exclusively by n.

reduced to around 0.14 as a result of free electron absorption within the skin depth. In terms of the wavenumber $k = k_1 - ik_2 = \omega\sqrt{\varepsilon}/c$, the skin depth ($\delta$) is defined as

$$\delta = \frac{1}{k_2} \equiv \frac{c}{\omega} \left(\frac{2}{\sqrt{{\varepsilon_1}^2 + {\varepsilon_2}^2} - \varepsilon_1} \right)^{1/2} \quad (17)$$

which can be calculated from the Drude forms for ε_1 and ε_2.

There are two limiting cases of the skin depth with practical consequences depending on the frequency and the properties of the conducting material, as illustrated below.

Case 1

$$\omega < \omega_p, \ \omega\tau \ll 1, \ \varepsilon_2 \approx \frac{{\omega_N}^2 \tau}{\omega} \gg |\varepsilon_1|, \ \delta \approx \left(\frac{2}{\omega\mu_0\sigma} \right)^{1/2}$$

Case 2

$$\omega > \omega_p, \ \omega\tau \gg 1, \ \varepsilon_2 \approx \frac{{\omega_N}^2}{\omega^3\tau} \ll \varepsilon_1, \ \delta \approx \frac{2m^*}{Z_\infty e^2} \frac{\omega^2\tau}{n}$$

(assuming $\omega < E_g/\hbar$, $\varepsilon_1 \approx \varepsilon_\infty$, $Z_\infty \approx 377\Omega/\sqrt{\varepsilon_\infty}$)

Case 1 applies to ITO and related materials well below the plasma frequency (*e.g.* for incident microwave radiation, where typically $\omega \approx 10^{10}$–10^{11} s^{-1}) and in this limit ITO is highly reflecting with very low transparency. Case 2 applies to ITO at visible wavelengths for electron densities up to around $n_{max} \approx 1.5 \times 10^{21}$ cm^{-3} and mobilities $\mu_e > 10$ cm^2 V^{-1} s^{-1}. Hence, case 2 is the most important limit when assessing the transparency of ITO layers for visible light.

In case 2 we observe that approximately $\delta \propto \omega^2\mu_e/n$ (assuming that $\tau_0 \propto \tau$), so that increasing n leads to the expected reduction in δ as ω_p is reduced, together with an associated reduction in the transparency of the film. However, an *increase* in mobility for fixed n in this limit actually *increases* the skin depth and *increases* the film transparency. An exact calculation of the skin depth for an incident free-space wavelength of 800 nm is shown in Fig. 8, where case 2 is attained towards the bottom right hand portion of this graph. Since approximately $\delta \propto \omega^2$, the smallest value of δ within the visible spectrum occurs for red light. Hence, if the transparency of an ITO film is deemed adequate for red light, it will be even better for blue light since then we are even further from the plasma edge and the skin depth is approximately four times larger.

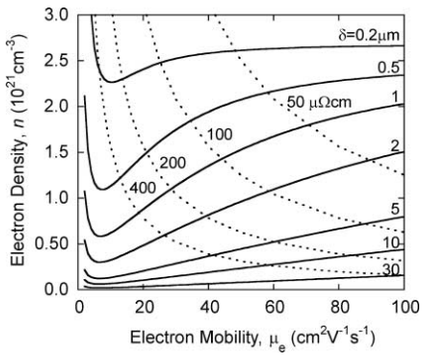

Fig. 8 Contours of the skin depth δ (in units of μm) as a function of electron density n and mobility μ_e for an ITO film at an incident wavelength $\lambda = 800$ nm. Also shown are contours of constant resistivity $\rho = 1/ne\mu_e$ in units of μΩ cm (dotted lines). The simple formula for the skin depth $\delta \propto \omega^2\tau/n \propto \omega^2\mu_e/n$ is valid towards the bottom right hand corner of this graph.

This result has an important consequence if high transparency films with very low sheet resistance $R_s = 1/ne\mu_e t$ are required. Immediately we see from Fig. 8 that this is best achieved by increasing the mobility in preference to n, since increasing n leads to a proportionate reduction in the skin depth as the plasma frequency is increased. For example, if μ_e could be doubled whilst preserving high n then, assuming that $\mu_e > 10$ cm^2 V^{-1} s^{-1} to start with, in principle a film could be grown to twice the thickness whilst preserving the same transparency; hence, R_s for this film could be reduced by a factor of four, which is a very significant result that illustrates the importance of maximising μ_e. This might be achieved by improved film processing so that μ_e approaches its intrinsic limit, which in ITO is set at around 100 cm^2 V^{-1} s^{-1} as a result of ionised impurity scattering.[23]

5 Concluding remarks

We have reviewed the electronic structure of transparent conducting oxides. A simple effective mass model captures most of the essential materials physics of the isolated donor states. The composition-induced transition to the metallic, degenerate-electron state at electron densities of *ca.* 10^{19} cm^{-3} is discussed in terms of the Mott transition to the metallic state and the concept of minimum metallic conductivity. Above this critical (Mott) density Sn impurities are simply ionized and associated itinerant electrons occupy the bottom of the conduction band in the form of an electron gas. As noted elsewhere,[23] Sn ions behave as point scatterers of the itinerant electrons and such scattering events probably set the intrinsic limits for electron (carrier) mobilities. We have used the simple Drude model to investigate the parameters important for electromagnetic absorption in ITO films of high carrier density. A simple approximation for the skin depth above the plasma frequency shows that it is proportional to electron mobility. We have then used this result to show that an increase in mobility gives rise to an increase in the optical transparency of an ITO film associated with increased skin depth, and consequently reduced electromagnetic absorption. These results have bearing on the optoelectronic applications of alternative transparent conductor systems, where it may be possible to achieve even higher mobilities (*e.g.* $InGaO_3(ZnO)$, where values as high as 140 cm^2 V^{-1} s^{-1} have been reported[9]), provided that this is coupled with moderately high electron densities.

We hope that this quantitative approach allows one to gain insights into the fundamental materials physical limits (the performance limits) of these thin film materials. It now appears that such performance limits are now regularly being approached in known materials[23] and this fact sets the scene for major activities in the materials chemistry of new systems to optimize/enhance properties. In essence, enhanced performance must be achieved by increasing the carrier mobilities rather than the carrier density (equivalent to increasing the resistivity scattering time). As noted elsewhere,[6,14,24] the key to high carrier mobility is the presence of a broadly-dispersed (wide) free-electron s-like conduction band (Fig. 1) consistent with a large hopping (transfer) integral between neighbouring sites. The presence of a broad conduction band, and a low carrier effective mass reflect the need for substantial metal–metal contacts in the materials crystal structure. An s-like conduction band represents a uniform electronic charge density distribution and concomitant low scattering. In addition, Mryasov and Freeman[14] note that a simple, isotopic cubic crystal structure should, *a priori*, afford fewer ionized and neutral scattering centres than more complex ITO-type structures having non-equivalent metal and O sites, plus several types of vacancies and disorder.

Such considerations – encompassing both the materials physics and materials chemistry – offer great potential for broaching the current performance limits on maximum mobility and transparency imposed by the crystal and electronic structure.

Acknowledgements

The authors wish to thank the EPSRC and Merck limited for funding and Ms Katie Johnson for help in assembling the references for the manuscript.

References

1 J. I. Hamberg and C. G. Granqvist, *J. Appl. Phys.*, 1986, **60**, R123.
2 C. G. Granqvist, *Appl. Phys. A*, 1991, **52**(2), 83.
3 C. G. Granqvist and A. Hultaker, *Thin Solid Films*, 2002, **411**, 1.
4 J. M. Phillips, R. J. Cava, G. A. Thomas, S. A. Carter, J. Kwo, T. Siegrist, J. J. Krajewski, J. H. Marshall, W. F. Peck, Jr. and D. H. Rapkine, *Appl. Phys. Lett.*, 1995, **67**, 2246.
5 B. G. Lewis and D. C. Paine, *MRS Bull.*, 2000(August), 22.
6 A. J. Freeman, K. R. Poeppelmeier, T. O. Mason, R. P. H. Chang and T. J. Marks, *MRS Bull.*, 2000(August), 45.
7 R. Wang, L. L. H. King and A. W. Sleight, *J. Mater. Res.*, 1996, **11**(7), 1659.
8 C. Ktlic and A. Zunger, *Phys. Rev. Lett.*, 2002, **88**(9), 95501.
9 K. Nomura, H. Ohta, K. Ueda, T. Kamiya, M. Hirano and H. Hosono, *Science*, 2003, **300**, 1269.
10 A. Porch, D. V. Morgan, R. M. Perks, M. O. Jones and P. P. Edwards, *J. Appl. Phys.*, 2004, **95**(9), 4734.
11 A. Porch, D. V. Morgan, R. M. Perks, M. O. Jones and P. P. Edwards, *J. Appl. Phys.*, 2004, 043419JAP; to be published October 2004.
12 I. Tanaka, M. Mizuno and H. Adachi, *Phys. Rev. B.*, 1997, **56**, 3536.
13 J. C. C. Fan and J. B. Goodenough, *J. Appl. Phys.*, 1977, **48**, 3524.
14 For more recent developments see: O. N. Mryasov and A. J. Freeman, *Phys. Rev. B*, 2001, **64**, 233111.
15 W. Kohn, *Solid State Physics V*, Academic Press, Inc., New York, 1957, p. 257.
16 N. F. Mott, in *Metal–Insulator–Transition*, Taylor and Francis, London, 1990.
17 (*a*) P. P. Edwards and C. N. R. Rao, in *The Metallic and Nonmetallic States of Matter*, Taylor and Francis, London, 1984; (*b*) P. P. Edwards, in *The New Chemistry*, ed. N. Hall, Cambridge University Press, 2000, pp. 85–114.
18 J. A. Marley and R. C. Dockerty, *Phys. Rev. A*, 1965, **304**, 140.
19 M. R. Graham, C. J. Adkins, H. Behar and R. Rosenbaum, *J. Phys.: Condens. Matter*, 1998, **10**, 809.
20 A. I. Ioffe and A. R. Regel, *Progress in Semiconductors*, 4, ed. A. F. Gibson. F. A. Kroger and R. E. Burgess, Heywood, London, 1960, p. 237.

21 See, for example: H. Fritzche, in *The Metal–Monmetal Transition in Disordered Systems*, ed. L. R. Friedman and D. P. Tunstall, Proceedings of the Nineteenth Scottish Universities Summer School in Physics, Edinburgh, The School, 1978, pp. 193–238.
22 P. P. Edwards and M. J. Sienko, *Phys. Rev. B*, 1978, **17**, 2575.
23 J. R. Bellingham, W. A. Phillips and C. J. Adkins, *J. Mater. Sci. Lett.*, 1992, **11**, 263.
24 D. R. Kammler, T. O. Mason, D. L. Young, T. J. Coutts, D. Ko, K. R. Poeppelmeier and D. L. Williamson, *J. Appl. Phys.*, 2001, **90**, 5979.