The shallow-to-deep instability of hydrogen and muonium in II–VI and III–V semiconductors

2003 J. Phys.: Condens. Matter 15 R1727

(http://iopscience.iop.org/0953-8984/15/46/R01)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **15** (2003) R1727–R1780 PII: S0953-8984(03)39533-5

TOPICAL REVIEW

The shallow-to-deep instability of hydrogen and muonium in II–VI and III–V semiconductors

S F J Cox

ISIS Pulsed Muon Facility, Rutherford Appleton Laboratory, Chilton, Oxfordshire OX11 0QX, UK
and
Department of Physics and Astronomy, University College London, WC1E 6BT, UK

Received 12 June 2003
Published 7 November 2003
Online at stacks.iop.org/JPhysCM/15/R1727

Abstract

The structure and electrical activity of monatomic hydrogen defect centres are inferred from the spectroscopy and charge-state transitions of muonium, the light pseudo-isotope of hydrogen. Introductions are given to all these topics. Special attention is paid to the shallow-donor behaviour recently established in a number of II–VI compounds and one III nitride. This contrasts with trapped-atom states suggestive of an acceptor function in other members of the II–VI family as well as with the deep-level amphoteric behaviour which has long been known in the elemental group-IV semiconductors and certain III–V compounds. The systematics of this remarkable shallow-to-deep instability are examined in terms of simple chemical considerations, as well as current theoretical and computational models. The muonium data appear to confirm predictions that the switch from shallow to deep behaviour is governed primarily by the depth of the conduction-band minimum below the vacuum continuum. The threshold electron affinity is around 3.5 eV, which compares favourably with computational estimates of a so-called pinning level for hydrogen (+/−) charge-state transitions of between −3 and −4.5 eV. A purely ionic model gives some intuitive understanding of this behaviour as well as the invariance of the threshold. Another current description applies equally to covalent materials and relates the threshold to the origin of the electrochemical scale. At the present level of approximation, zero-point energy corrections to the transition levels are small, so that muonium data should provide a reliable guide to the behaviour of hydrogen. Muonium spectroscopy proves to be more sensitive to the (0/+) donor level than to the (+/−) pinning level but, as a tool which does not rely on favourable hydrogen solubility, it looks set to test further predictions of these models in a large number of other materials, notably oxides. Certain candidate thin-film insulators and high-permittivity gate dielectrics appear to be uncomfortably close to conditions in which hydrogen impurity may cause electronic conduction.

This review is dedicated to the memory of T L Estle (1931–2002).

0953-8984/03/461727+54$30.00

Contents

1. Context: the electrical activity of interstitial hydrogen

1.1. Hydrogen impurity

This review concerns hydrogen as a largely unavoidable impurity in semiconducting material, influencing both fabrication processes and final electrical properties. It focuses on the isolated defect centres, H^+, H^0 and H^-—respectively the interstitial proton, the trapped atom or its derivative neutral centres, and the hydride ion where appropriate. Such states normally account only for a minute proportion of incorporated hydrogen—the majority is either bonded, paired with other defects or impurities in passivation complexes, or present as interstitial molecular hydrogen. As an example, hydrogen levels can reach several atomic per cent in ZnO and CdS layers in certain designs of solar cell, as illustrated in figure 1. Most of this is quite inert or actually improves electronic performance by passivating unwanted defects. This is exemplified by the case of amorphous silicon—another photovoltaic material—where deliberate hydrogenation is used to saturate dangling bonds, removing their electrically active levels from the energy gap (Street 1991). Nonetheless, the sparse monatomic hydrogen centres—those that are not paired with other defects or impurities—are now understood to have an electrical activity of their own: that is, they themselves introduce defect levels into the energy gap and participate in the trapping, release and scattering of charge carriers. They have a disproportionate importance in this respect, therefore, and increasingly so at high temperatures, where passivation complexes become dissociated. In silicon, the behaviour is at last reasonably well understood in terms of a fascinating interplay of charge state and crystallographic site. The task at hand is to examine how well this understanding carries over to other classes of semiconductor, particularly the wide-gap materials that are now on the verge of important applications, notably in optoelectronics.

1.2. Charge states; defects and transition levels; negative U

It is easy to see that H^+ will be stabilized in cation vacancies, for example, or H^- in anion vacancies, but we are concerned here with interstitial rather than substitutional sites. In this regard, the possibility that hydrogen can exist in all three charge states, both in covalent and ionic semiconductors, is rather remarkable. It relies on several circumstances. One is that the electronic disturbance is localized within the screening length, so that distinct charge states are meaningful[1]. Another is that the two transition points both lie deep in the energy gap, as

[1] For instance, the screening length is about 3 nm in highly doped n-type Si (Chow *et al* 2000).

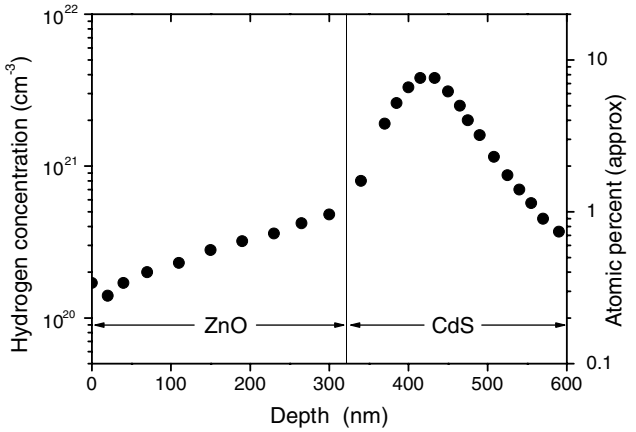

Figure 1. Hydrogen concentration profiles through the ZnO window and CdS buffer layers of a solar cell, determined by nuclear reaction analysis (Krauser *et al* 2000). In these sputtered films, hydrogen concentration exceeds its solubility limit in the bulk materials. The ZnO layer is naturally n-type (probably due at least in part to the hydrogen impurity itself) and completes the solar cell heterojunction.

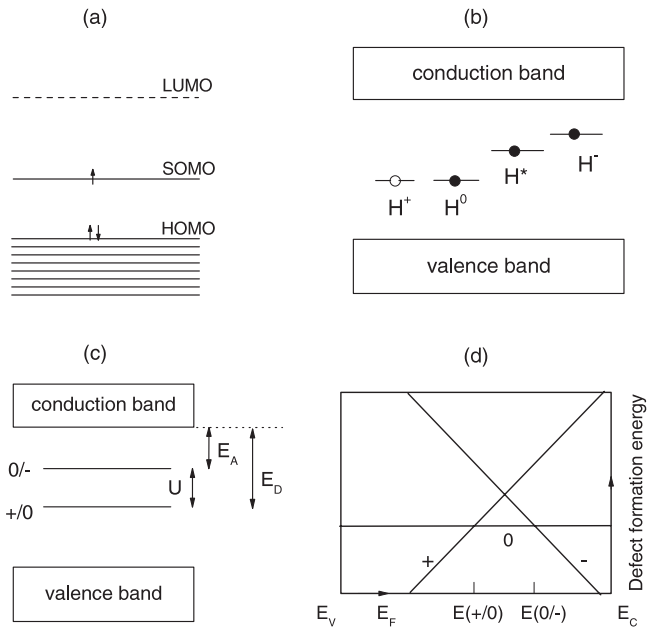

Figure 2. Different representations of amphoteric hydrogen centres, for the positive-U case. Only the neutral centre is shown in (a), with the usual chemical notation: HOMO = highest (doubly) occupied molecular orbital, SOMO = singly occupied molecular orbital and LUMO = lowest unoccupied molecular orbital. The other charge states are depicted in (b)–(d). Also depicted in (b) is the possibility of a metastable neutral centre (H*) which plays no role in thermodynamic equilibrium, where the transition points in (c) correspond to crossings of the total energies in (d).

was generally supposed hitherto. This is illustrated in figures 2(a)–(d) for the case where the Hubbard or Anderson U is positive: this quantity is then essentially the Coulomb repulsion

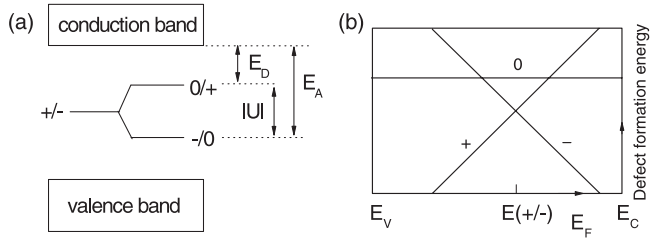

Figure 3. Inverted ordering of the (0/+) and (−/0) transition points for the case of negative U (compare figures 2(c) and (d)). H^0 is never thermodynamically stable; instead, the charge state switches from H^+ in sufficiently p-type material to H^- in sufficiently n-type. For a wide range of conditions, E_F is pinned at the (+/−) level.

energy incurred on going from the singly to the doubly occupied defect—from H^0 to H^-, in our case. If these are the only defects present, and there is one electron per defect, they adopt their neutral paramagnetic (H^0) state. In the presence of other impurities, notably shallow dopants, each of the three charge states can be stabilized in turn by an appropriate Fermi energy, E_F. Thus in p-type material, only H^+ is stable, all deep levels being emptied by a low Fermi energy. The charge state switches from H^+ to H^0 and from H^0 to H^- as E_F is notionally raised through the (0/+) and (−/0) thresholds.

In fact, hydrogen in semiconductors may constitute a negative-U system, so that the (0/+) and (−/0) levels are inverted, as in figure 3. The Coulomb repulsion energy in the H^- ion is then more than offset by a change of configuration or lattice energy. This is readily possible for an interstitial defect, since the different charge states can occupy distinct crystallographic sites as well as inducing different lattice relaxations. An eminently readable account of this behaviour is given by Watkins (1981) for the case of interstitial boron in Si. There is now a general, but by no means unanimous, consensus that interstitial hydrogen behaves in this manner in conventional semiconductors such as Si and GaAs, as explained in appendix A.

1.3. The pinning level

For negative-U defects, the neutral centres are never thermodynamically stable. If they are the only defects present, they disproportionate into the positive and negative ions:

$$2H^0 \rightarrow H^+ + H^-. \tag{1}$$

The neutral centres may exist transiently, of course, and it is worth mentioning that they are recovered under conditions of bandgap illumination (as is indeed the case for hydrogen in silicon). Otherwise, their equilibrium charge state is again determined by the relative concentration of shallow dopants. This is clear in the usual type of representation (figure 3(b)), although the implication that E_F can be varied continuously through the gap is somewhat misleading. When the negative-U centres are in abundance, the Fermi energy is in fact pinned at the (+/−) transition level for a wide range of extrinsic electron concentrations (i.e. until all the negative-U centres are doubly occupied: see, for example, Adler and Yoffa (1976)). This pinning is robust against wide variations of temperature and, for large concentrations of the negative-U centres, makes the material virtually immune to doping.

Only when the pinning level falls at or above the conduction-band (CB) edge can negative-U defects *cause* conductivity under equilibrium conditions, rather than oppose it, as we see in the concluding sections.

It is useful to identify the (0/+) transition with the behaviour of a donor and the (0/−) transition with that of an acceptor: the definitions are sometimes debatable but the electronic wavefunctions involved justify this identification for H in Si and GaAs, for example. Referring to the diagrams of the defect total energies as a function of E_F in figures 2(d) and 3(b) (some authors refer to these as the *formation energies* for each charge state) the slope of each line is equal to the nominal charge, so that the pinning level is given by

$$E(+/-) = \tfrac{1}{2}\{E(0/+) + E(-/0)\} \tag{2}$$

and the Hubbard or Anderson U by

$$U = E(0/+) - E(-/0). \tag{3}$$

1.4. Lifetimes and out-of-equilibrium conditions

The putative interstitial hydride ion H^- should only be stable in heavily doped n-type material, where the high E_F favours double occupation of deep levels. As well as the question of screening, high doping levels raise the related issue of the lifetime of the different charge states. To define a neutral state, e.g. a trapped-atom H^0 state, the lifetime must greatly exceed the classical period of the Bohr orbit—a condition which is readily satisfied in intrinsic semiconductors at low temperatures. In magnetic resonance spectroscopy, the condition becomes more stringent: for a paramagnetic state such as H^0 to be recognizable, e.g. via resolved hyperfine splittings in an ESR spectrum, the spin-state lifetime of the unpaired electron must exceed many periods of the hyperfine oscillation (or, in high field, of the electron Larmor precession, whichever is the faster: see appendix B). This condition may not be met in an n-type sample, where fast spin exchange with the conduction electrons would make a paramagnetic H^0 centre indistinguishable from an electronically diamagnetic centre such as the H^- ion (see, for example, Chow *et al* 2000, Senba 2001). The criterion for actual capture of the second electron, making the ionic configuration $1s^2$, may not be precisely known, but it is customary to assume that H^- can exist at sufficiently high doping levels, both in polar semiconductors such as ZnSe and valence systems such as Si or GaAs. This is despite its large radius which, at the value usually given as 208 pm, would be hard to accommodate in most interstitial locations; the value of 140 pm, typical of H^- in metal hydrides, is perhaps more appropriate.

In highly ionic materials, including some oxides, high concentrations of extrinsic carriers will be difficult to achieve, so that the concept of thermodynamic equilibrium implied by figures 2(d) and 3(b) does not apply: for most practical purposes, the exchange of electrons with a notional Fermi sea simply cannot take effect fast enough. In fact, out-of-equilibrium states and processes will be important in all classes of semiconductor. Many electronic processes come into this category, with the result that specific electrical or spectroscopic measurements may show switching levels that differ from the thermodynamic transition points. As well as the response to illumination—evidently important to applications in photovoltaics or optoelectronics—conditions of radiolysis or ion implantation are relevant to the present work. Other cases in point concern the (+/−) and (−/0) transitions. The former must proceed via a neutral state, which will have a finite lifetime (during which the site change presumably occurs). The latter corresponds to ionization of H^-, which may leave H^0 in a crystallographic site that is not its most stable; the ionization energy without site change may be the spectroscopically accessible parameter but it does not define the thermodynamic acceptor level (see appendix A).

1.5. Compensation versus doping

The dual function of donor and acceptor is known as amphoteric (sometimes ambivalent) behaviour. Whether U is positive or negative, a consequence of the transition points lying

deep in the energy gap is that hydrogen releases electrons when E_F is low, i.e. loses them to remote acceptors in p-type material, but captures electrons when E_F is high, attracting them from remote donors in n-type material. As an impurity, therefore, hydrogen acts as a compensating defect and reduces the conductivity created by deliberate shallow dopants—as indeed it does in its more familiar, though quite distinct, role as a passivation agent.

In this context, the discovery in CdS by Gil *et al* (1999) of a muonium centre in which the electron was unusually weakly bound came as something of a surprise. (Here muonium may be regarded as the radioactive light isotope of hydrogen, as explained in section 2.) The observation implied that the analogous hydrogen centre would have all the characteristics of a *shallow* donor. Independent theoretical work was in progress at that time, however, resulting in predictions of hydrogen as a possible source of conductivity in another II–VI material, namely ZnO (Van de Walle 2000), as well as in more complex oxides (Park and Chadi 1999, 2000). In fact, hydrogen impurity had been implicated in the puzzling n-type conductivity of ZnO much earlier (e.g. Thomas and Lander 1956, Hutson 1957) but either this work had been subsequently overlooked or it had not proved that hydrogen was the actual dopant rather than the agent which activated some other defect or impurity. A similar story in InN is related below. The paper by Van de Walle (2000) makes a clear exposition of the contrast between the deep-level amphoteric behaviour of H in Si, where it always opposes the prevailing conductivity as a compensating defect, and its role in ZnO as the likely *cause* of the prevailing n-type conductivity. The prediction was quickly confirmed by spectroscopic characterization of the undissociated donor, i.e. of the paramagnetic state, using both the muonium isotope (Cox *et al* 2001a) and hydrogen itself, i.e. protium[2] (Hofmann *et al* 2002). Meanwhile, shallow muonium states were also detected and characterized in CdSe and CdTe so that, by inference, hydrogen could act as a dopant in its own right in at least four of the II–VI compounds.

1.6. Plan of the article

The results of this initial survey have already been given by Gil *et al* (2001). In the present review, the methodology is described and evaluated at greater length, with a brief introduction in section 2 and with other aspects explained in later sections as required. The survey of shallow-donor muonium states is updated and extended to the III–Vs, including nitrides, and the systematics interpreted in terms of the nature of the deep-to-shallow instability in succeeding sections. The predictions of the current models for electronic (as opposed to protonic) conduction induced by hydrogen impurity in oxides are examined in the concluding sections, as is the extent to which muonium data can be expected to carry over as a guide to the electrical activity of protium.

Appendices include reference material. The deep-level amphoteric behaviour of hydrogen which has hitherto been assumed to be the norm is described in appendix A, exemplified by the interplay of site and charge state for muonium in Si and GaAs. Its precise relation to the behaviour of protium is still not without controversy. Appendix B gives some details of the muonium hyperfine spectroscopy and examples of different variants of the experimental methods using transverse, longitudinal and even zero magnetic field. Appendix C gives a discussion of the relation between ionization energy, donor depth and electron capture mechanisms, as appropriate to muon implantation and muonium formation. Appendix D is a compendium of muonium hyperfine constants.

[2] We use this somewhat pedantic chemical notation where it is useful to distinguish muonium and protium as two isotopes of hydrogen.

2. The muon–proton and muonium–protium analogies

2.1. μSR *spectroscopy*

For a better understanding of electrical activity, atomistic pictures of the defect centre in its individual charge states are essential. These require knowledge of the electronic structures and—for an interstitial defect—the crystallographic site or sites occupied. This sort of information has been extremely hard to obtain for isolated hydrogen centres, to the extent that much of our present understanding comes in fact from spectroscopic studies, not of the hydrogen centres themselves, but of their muonium counterparts. The pioneering studies of muonium in Si and GaAs around the 1970s and 1980s were particularly influential and recent studies of the more ionic III nitrides may prove equally so. Meanwhile the techniques of computational chemistry have advanced enormously, themselves providing the atomistic pictures that were previously lacking, but the muonium data continue to provide stringent quantitative tests of the predictions.

The shallow-state studies reviewed here are amongst the most recent examples of the use of muonium as a pseudo-isotope and model for hydrogen. The methodology relies on positive muons (the elementary particles common in cosmic rays and usually denoted μ^+) mimicking protons faithfully in their solid-state chemistry. If this analogy seems somewhat exotic, not to say contrived, it is the unique properties of muon production and decay which form the basis of a spectroscopy whose sensitivity, selectivity and timescale are all hard to match with more conventional methods. This is μSR spectroscopy, standing for muon spin rotation, muon spin relaxation or muon spin resonance according to the precise variant of the technique (Brewer *et al* 1975); the acronym was clearly contrived to resemble ESR (electron spin resonance), i.e. to emphasize similarities with magnetic resonance. In fact, μSR spectroscopy is equally sensitive to electronically paramagnetic and diamagnetic muonium defect centres, so that there are also links and analogies with NMR (nuclear magnetic resonance) and with ENDOR (electron nuclear double resonance—see appendix B).

2.2. *Muon facilities*

While cosmic-ray muons could in principle be used, given huge samples and considerable patience[3], routine use of the spectroscopy in practice relies on the availability of spin-polarized beams of positive muons at certain accelerator laboratories. There are unfortunately no portable or bench-top sources. Four such facilities are currently operational, those at the Paul Scherrer Institute near Zurich and the TRIUMF Laboratory in Vancouver providing intense continuous beams of these particles and those at the ISIS Facility near Oxford and KEK near Tokyo providing beams which are less intense on average but which have certain advantages in being pulsed (see, for example, Cox 1987). Each spectrum and data point in the figures below typically represents no more than 1 h of beamtime at these modern sources (although the demonstration spectrum of figure 6 was run for rather longer). Several other muon sources in the USA and former Soviet Union were also historically important to semiconductor studies, as documented in an important review by Patterson (1988). Work up to the late 1980s covers much of what is known of the deep muonium states, including the trapped-atom states in II–VI materials such as MgO, ZnS and ZnSe. As a historical perspective, the ISIS facility came on line in 1987, around the time that attention was turning from the purely spectroscopic characterization of muonium to a study of its charge-state transitions as a model for electrical

[3] As you read this review, muons produced by primary interactions in the upper atmosphere are raining down on you at a steady rate of about one a second! Some undergraduate teaching laboratories run muon-lifetime and demonstration muon spin rotation experiments using stopped cosmic rays.

activity. Patterson's (1988) article as well as subsequent reviews (see e.g. Kiefl and Estle 1992, Chow *et al* 1998, Lichti 1999) describe the principles and techniques of the spectroscopy comprehensively. A new review by Lichti (2003a) summarizes the latest applications to the III–V semiconductors, including nitrides. Nonetheless the essentials are reproduced below for convenience.

2.3. Muon implantation

All the spectroscopy described in this review uses 4 MeV beams of positive muons, stopped in the material of interest. At this energy, the muons typically penetrate to a few tenths of a millimetre, so that thick-film or bulk semiconductor samples are needed. (4 MeV is the lowest energy at which muon beams are easily produced, set by the decay energy of the parent pions; moderation techniques enabling surface or thin-film studies are still under development and have yet to be applied systematically to semiconductors: see, for example, Morenzoni *et al* (2003).) As soon as they have lost their incoming kinetic energy and become thermalized, the implanted muons are assumed to adopt the same positions in the host lattice as would interstitial protons. These two elementary particles have the same unit positive charge and, although the muon is an order or magnitude lighter than the proton, it generally localizes in equivalent potential wells, most importantly at interstitial sites[4]. Differences between their properties which are fundamental in high-energy or particle physics are largely irrelevant to their chemical behaviour at thermal or even epithermal energies. Thermalization does not here imply thermodynamic equilibrium and we shall see immediately that population of metastable states, both in terms of crystallographic site and electronic charge state, have been particularly illuminating: this is a consequence of the microsecond timescale of μSR spectroscopy, set by the muon lifetime of just over two millionths of a second, $\tau_\mu = 2.2\ \mu$s.

2.4. Muonium: a radioactive light isotope of hydrogen

If the muon binds an electron, the resultant hydrogen-like atom is effectively a light isotope of hydrogen. This atomic state is known as muonium and given the chemical symbol Mu. It is hydrogen in which the proton has, so to speak, been replaced by the positive muon: $\mathrm{Mu} = [\mu^+\mathrm{e}^-]$[Note 5]. The reduced electron mass is the same as in protium to better than 0.5%. In their vacuum states, therefore, muonium and protium have essentially the same Bohr radius and ionization potential. Tables 1–3 include relevant parameters and properties. This close analogy between muonium and protium is assumed also to apply to their solid state chemistry, i.e. remains valid on going from the free atoms to their derivative defect centres in solids.

2.5. Interplay of site and charge-state

Atomic-like states of hydrogen are well known in oxides and halides (the traditional hosts for ESR studies), their hyperfine constants differing slightly from that of the free atom and varying from one material to another, depending on the degree of overlap or admixture of the 1s(H) wavefunction with the host atoms. Analogous muonium states are well known and their hyperfine systematics show exactly the same pattern (Spaeth 1986), subject to a small zero-point energy correction. The charge-state transitions of hydrogen in these materials have, as yet, hardly begun to be examined but look set to become of some importance, especially

[4] This is quite different from the positron, 200 times lighter again, which generally adopts wavelike propagating states, localizing only at defect sites such as vacancies.

[5] Both the name and symbol have belatedly been adopted and approved by IUPAC. For a discussion of nomenclature and justification of isotope status, see Walker (1983).

Table 1. Chemical and magnetic properties of the free particles. (We reserve the symbols μ^+ and p for the initially energetic incoming particles, distinguishing the thermalized positive ions by the chemical notation H^+ and Mu^+.)

Particle	Symbol	Mass (m_e)	Mass (m_p)	Spin	Gyromagnetic ratio γ
Electron	e	1	5×10^{-4}	1/2	$2\pi \times 28$ MHz mT^{-1}
Muon	μ^+	207	$0.113 \approx 1/9$	1/2	$2\pi \times 136$ kHz mT^{-1}
Pion	π^+	273	0.149	0	
Proton	p	1836	1	1/2	$2\pi \times 42.6$ kHz mT^{-1}

Table 2. Decay modes of the particles. That of the pion, generated when nuclei of a light element such as carbon are bombarded with high-energy protons, is the basis of spin-polarized muon production. That of the muon is the basis of the detection method and polarization analysis in μSR spectroscopy (only the positrons are detected). The negative muon plays the role of a heavy electron, orbiting close to atomic nuclei. As a result, negative muons can be captured prematurely by atomic nuclei. Whereas certain applications simulating acceptors in semiconductors are noteworthy (see e.g. Mamedov *et al* 2000) they receive no further mention in this review. The positive muon plays the role of a light proton and is the focus of our attention; its lifetime is independent of its chemical environment.

Particle	Symbol	Lifetime	Decay mode	Antiparticle
Electron	e	∞		Positron (e^+)
Muon	μ^+	2.2 μs	$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$	Negative muon (μ^-)
Pion	π^+	26 ns	$\pi^+ \rightarrow \mu^+ + \nu_\mu$	Negative pion (π^-)
Proton	p	∞		Antiproton ($\bar{p}$)

Table 3. Free-atom properties: muonium as a light isotope of hydrogen.

Isotope	Chemical symbol	Mass (m_e)	Reduced mass (m_e)	Bohr radius a^* (nm)	Ionization energy $R_y = I(0/+)$ (eV)
Protium	^{1}H	1837	0.9995	0.0529	13.60
Muonium	Mu	208	0.9952	0.0532	13.54

in oxides. The existence of muonium states with similar trapped-atom character, formed on muon implantation into MgO, ZnS and ZnSe (Spencer *et al* 1983, Patterson 1988), is especially relevant to the present survey of the II–VI compounds.

The term 'trapped atoms' is used here to imply configurations most resembling free atomic hydrogen or muonium, where the highest spin density, i.e. the greatest probability density for the unpaired electron, remains centred on the proton or muon. Their localization within interstitial cages does not imply immobility and indeed such light impurity atoms often diffuse readily through the lattice, with mechanisms on the verge of quantum diffusion. It is the largest interstitial cages that are occupied, in order to minimize confinement energy; there is minimal elastic distortion and no directional bonding. It is curious that ESR studies of atomic hydrogen have not been reported for the tetrahedrally coordinated materials, even though the interstitial cavities are generally larger than in octohedrally coordinated structures. By contrast, μSR studies of such materials have been particularly fruitful. Muonium is found to lose rather more of its atomic character in semiconductors than in the octahedrally coordinated oxides and halides, with 1s spin density on the muon reduced to about 50% for Mu^0 for Si, for example. (This variation is shown in figure 22 below, where we return to this question; a comprehensive compendium of values is given in appendix D.) The failure of ESR to detect the trapped-atom H^0 state in Si is almost certainly due to its high mobility and propensity to

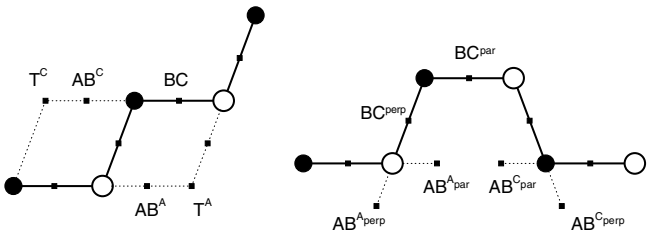

Figure 4. Principal interstitial sites for hydrogen and muonium in tetrahedrally coordinated lattices: (a) zincblende or diamond type (cubic) and (b) wurtzite (hexagonal). T = tetrahedral cage centre, BC = bond centre and AB = antibonding. Superscripts denote sites adjacent to the cation (C) or anion (A) and distinguish inequivalent sites where necessary. The BC site is unique in (a) but not in (b); in both structures its occupation results in considerable local distortion.

pair with other defects or impurities. The microsecond timescale of μSR spectroscopy can be crucial in characterizing such highly mobile or reactive states.

That the trapped atom is not the only, nor even the most important, state of isolated hydrogen in semiconductors became apparent largely as a result of muonium spectroscopy. In silicon, the atomic Mu^0 state proves to be only metastable. It is also not the neutral state of the deep donor: the latter has a quite different electronic structure, associated with a bond-centred rather than cage-centred site. The interplay of site and charge state in tetrahedrally coordinated semiconductors with the diamond or zincblende structures is examined in appendix A. In brief, for the group-IV elements Si, Ge and diamond, and for the III–V compounds GaAs and GaP, just four states suffice. The positive ion adopts the bond-centre (BC) site, where valence-bond electron density is maximal, and is denoted Mu^+_{BC}; the negative ion is repelled to the cage centre known as the T-site, where valence charge is minimal, thus Mu^-_T. These sites are illustrated in figure 4(a). The neutral centres may exhibit metastability, with a fine energy balance between the T and BC sites, Mu^0_T and Mu^0_{BC}. The analogous nomenclature and similar interplay of the four states H^+_{BC}, H^0_{BC}, H^0_T and H^-_T is now commonly adopted in discussions of hydrogen diffusion and charge-state transitions in silicon (see, for example, Bonde Nielsen *et al* 1999).

The manner in which these site preferences evolve in more ionic semiconductors, and on going from the zincblende to the wurtzite structure (see figure 4(b)), has already been examined for muonium in the group-III nitrides (Cox *et al* 2002b, Lichti 2003a). It is noteworthy that, for AlN and GaN at least, there is as yet no evidence of occupation of BC sites, either by the neutral or positive centres. Mu^+ is found in close proximity to the nitrogen anion, as expected, and Mu^- adjacent to the cation, at respective antibonding (AB) sites displaced from the cage or channel centres. It is now of interest to see which sites are important for hydrogen impurity in the II–VI materials, some of which can exist in both zincblende and wurtzite structures.

2.6. Intrinsic versus extrinsic sites

It is worth emphasizing that, with few exceptions, muonium is confined to interstitial sites—a term which we use to designate any site in otherwise stoichiometric regions, including bond centres. In this review, we do not consider the multitude of defect sites or processes that can control the incorporation of hydrogen or water, even though these are especially important for protons in oxides (see for instance Catlow *et al* 1995). Neither do we consider reactions of atomic hydrogen which may proceed via a change of cation oxidation state (see the discussion of equation (18) below). We bear in mind that hydrogen which has been incorporated chemically, during deposition or processing of the material, is likely to be paired with other defects or

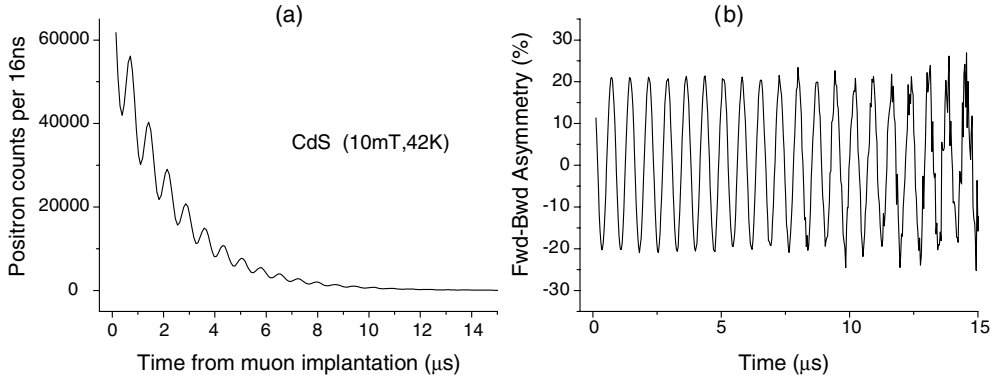

Figure 5. μSR histogram (a) and Larmor precession signal (b) for Mu^+ in CdS, recorded in 10 mT and at 42 K[Note 6]. The muon site is inferred from the orientation dependence of the (just discernible) damping rate (Gil *et al* 2001).

impurities. We encounter one example in the present survey—in ZnO—where the muon and proton sites may differ for this reason. Protons which are implanted from ion beams should for the most part reach the same interstitial sites as implanted muons (as for the case of the AA9 centre, i.e. H^0_{BC}) but they may also trap in vacancies created by the implantation process (see, for example, Gorelkinskii and Neninnyi 1987, 1991, Bech Nielsen *et al* 1997, Bech Nielsen and Grimmeiss 1989). The lighter muons create less radiation damage on implantation than protons and invariably thermalize downstream of any such damage (Brewer *et al* 1975). It is rare that they have time to seek out regions that are other than stoichiometric, on the microsecond timescale of μSR measurements. The yield of vacancy-trapped or oxygen-trapped muonium in silicon is minimal (Schefzik *et al* 2000a, 2000b) and it has been quite difficult to detect any signature of the muonium counterparts of passivation complexes, although an example has at last been confirmed in GaAs (Chow *et al* 1997, Chow 2003). It is true that the majority of μSR studies to date have been performed on good-quality material such as silicon and that defect sites may play a greater role in the newer semiconductors for which crystal growth procedures are still under development. For instance, it now appears that the exceptionally stable Mu^+ centre in AlN, reported by Cox *et al* (2002b) as visible up to 800 K, is only prominent in defective material!

If the sites occupied by hydrogen and muonium are the same, then the effects of the greater zero-point energy of the muon on local electronic structure, as sampled by spectroscopic parameters such as hyperfine or quadrupole interactions, are generally small (see, for example, Claxton and Cox 1994). The effect of nuclear zero-point energy on electrically active defect levels, both deep and shallow, is estimated in section 11.

3. The shallow muonium state in CdS

3.1. Muon spin rotation signals and muon site determination

Figure 5 shows Larmor spin precession for muons implanted into CdS at 42 K—the signal is typical also of a wide range of *higher* temperatures. It is detected via an intrinsic asymmetry

[6] ISIS-TC 2003: These data result from a demonstration experiment performed by students attending a μSR training course at ISIS. Loan of an Eagle–Picher single crystal of CdS by the Coimbra group is gratefully acknowledged; the precession signals were displayed and analysed using the WIMDA program due to FL Pratt; a maximum-entropy program from Southampton University was used to generate the frequency spectra of figures 6 and 7.

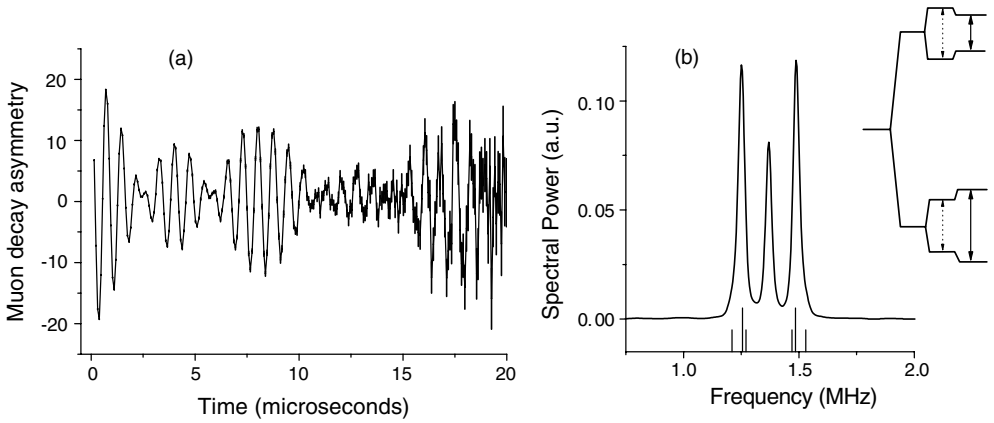

Figure 6. Detection of the undissociated muonium shallow donor, via beating in the μSR precession signal (a) and, equivalently, hyperfine splitting of the frequency spectrum (b). The inset shows the corresponding spin transitions between Mu^0 energy levels constructed by switching on in turn the electron Zeeman energy, muon Zeeman energy and hyperfine interaction (the three terms of equation (21) in appendix A); the dotted transition (not allowed for Mu^0) corresponds to the Mu^+ Larmor frequency, i.e. the central line. The magnetic field is applied perpendicular to the crystalline *c*-axis, an orientation chosen to minimize the effects of hyperfine anisotropy (compare simulated stick spectrum).

in the muon radioactive decay such that the β-emission is most intense in the instantaneous direction of the muon polarization. As each muon disintegrates, it emits a positron in a direction correlated with the parent muon's spin, and with sufficient energy to escape the sample and cryostat and be counted individually with scintillation detectors. Figure 5(a) is in fact the positron count-rate in a suitable array of detectors, binned in a histogram as a function of time elapsed since muon implantation: the underlying radioactive decay curve—$\exp(-t/\tau_\mu)$ with $\tau_\mu = 2.2$ μs—is seen to be modulated by the precession signal. The precession signal itself can either be extracted from single histograms of this sort (the usual procedure at continuous sources) or directly from the forward–backward asymmetry in count-rates between a pair of detectors or detector arrays (a convenient procedure for pulsed sources—see e.g. Cox 1987—and used here to generate figure 5(b)).

The oscillatory signal of figure 5(b) resembles a free induction decay signal, familiar in proton magnetic resonance; the precession frequency is in fact about three times higher than would be obtained for proton NMR at this field, the muon magnetic moment and gyromagnetic ratio being higher than those of the proton by this factor (table 1). Both are spin-1/2 particles, not susceptible to quadrupole interactions. The similarity belies very different methods in the preparation and analysis of the spin polarization, however. Proton NMR at such low fields and modest temperatures, relying on thermal-equilibrium populations of the spin states, would require at least 10^{17} spins for detection[7] and the character of the spectrum could then be dominated by proton–proton interactions. Figure 5, on the other hand, was accumulated by detecting just 10^7 muon decays and represents the extreme dilution limit. The two factors contributing to this high sensitivity *per spin* are the 100% polarization of the muon beam, largely retained on implantation, and the manner in which the spin precession is detected via single-particle counting at the enormously higher energy of the muon decay. Both are classic examples of parity violation (Brewer *et al* 1975).

[7] NMR spectra are usually recorded in a much higher field of several tesla. μSR sensitivity is not field dependent; in fact, direct measurements in zero-field are something of a speciality: see appendix B. This sensitivity is gained at the expense of resolution, which is limited in μSR by the muon lifetime.

At the temperature of figure 5, the muons are presumed to thermalize as the positive ion. We use the chemical nomenclature Mu^+, by analogy with H^+, reserving the particle physics nomenclature μ^+ for the free or energetic incoming particle. This is to emphasize the fact that the muon, like the proton, cannot remain free in condensed matter. It always seeks a region of high electron density to minimize its energy, moving to the BC site in covalent semiconductors such as Si, or even GaAs, for example, or by sticking to lone electron pairs on oxygen in certain oxides to form an analogue of the hydroxyl ion (see equation (8) below, plus discussion). The small damping of the μSR signal in CdS represents the sparsity of nuclear moments: only 24% of Cd nuclei in natural abundance carry spin and the corresponding percentage for S is quite negligible. Nevertheless the damping or relaxation rate is measurable and Gil *et al* (2001) conclude from its orientation dependence that the muon site is AB to the S anion. The calculated relaxation rates or linewidths, used for comparison with the data, did not allow for lattice distortion but the data appear to exclude the BC site, unless the bond extension is unexpectedly large.

3.2. Paramagnetic states: the ENDOR analogy

Figure 6 shows how a paramagnetic state is recognizable in μSR. This signal is for CdS cooled to 5 K, at which temperature some of the implanted muons pick up electrons in the course of thermalization to form Mu^0 centres. In place of the weakly damped signal of figure 5, strong beats have appeared in the muon spin rotation signal, indicative of several frequency components. In the frequency transform, three main lines are apparent. The central line again corresponds to muon Larmor precession, i.e. to those muons that still thermalize as Mu^+[Note 8]. The symmetrically placed satellite lines represent the Mu^0 centres that are paramagnetic by virtue of the single unpaired electron. Roughly speaking, the hyperfine field adds to or subtracts from the externally applied field, according to whether the electron is captured spin-up or spin-down. This is sketched in the inset to figure 6 and derived formally in appendix B. The hyperfine parameters are in fact so small in this CdS spectrum that 10 mT already corresponds to the high-field or Paschen–Back limit in which the muon–electron hyperfine interaction is effectively decoupled, i.e. dominated by the electron Zeeman energy! The two transitions therefore correspond to flips of the muon spin alone and are entirely analogous to the nuclear spin transitions observed in ENDOR spectroscopy. For the case of ZnO, a direct comparison of the μSR and proton-ENDOR spectra is now possible (see section 4 and appendix B).

3.3. Shallow-donor expectations

The remarkable feature here is the low value of the hyperfine splitting—just a few hundred kilohertz. For free muonium, the hyperfine constant is $A_0 \approx 4.5$ GHz, corresponding to the contact interaction with the electron in its unperturbed 1s atomic orbital[9]. For the trapped-atom states in ZnS and ZnSe it is about 3.5 GHz—some 80% of the free-atom value—indicating a degree of delocalization of spin density onto the host ions (Patterson 1988). In figure 6, the hyperfine splitting for Mu^0 centres in CdS is seen to be less than this by four orders of magnitude, suggesting very much greater extension of the electron wavefunction. There

[8] Additionally, this central line may subsume the signal from some muons landing outside the CdS sample on the metallic sample mount: the beam spot is comparable in size with the specimen, which are both about 1 cm^2 in this ISIS measurement.

[9] This is about three times the hyperfine frequency for protium, the ratio again being given roughly by the magnetic moment ratio for muon and proton, corrected by electron reduced masses. The precise ratio, known to eight significant figures, is used as a test of quantum electrodynamics and as a measure of the proton size. The muon, like the electron, is pointlike in all its interactions.

is a slight anisotropy, so it is the isotropic component or contact interaction which must be compared with the free-muonium hyperfine constant. These finer details of the spectroscopy are explained in appendix B; the order of magnitude is sufficient for the present purposes. The reduction from the free-ion value corresponds closely to the dilation of the effective Bohr radius, as required of an *effective-mass donor*[10], i.e. for an electron with its conduction-band effective mass m^*, sensing the bulk dielectric constant ε. The standard expressions for size and binding energy of effective-mass donors (see, for example, Stoneham 1975, Yu and Cardona 1996) are, in a common notation,

$$a^* = a_0\varepsilon/(m^*/m_e) \tag{4}$$

and

$$R^* = R_y(m^*/m_e)/\varepsilon^2. \tag{5}$$

Taking $\varepsilon = 9$ and $m^*/m_e = 0.2$ for CdS (Madelung 1996), the expected dilation factor of $a^*/a_0 = 45$ is to be compared with factor $\sqrt[3]{10^4} \sim 25$ from the Mu^0 hyperfine constant, assuming central spin density to vary as the inverse of the atomic volume. Likewise, the binding energy is expected to be reduced from that of free hydrogen or muonium (i.e. from the the Rydberg constant, $R_y = 13.6$ eV) to about 30 meV, this defining the depth of the shallow-donor level below the conduction band minimum (CBM). The experimental value is seen in the next section to lie in the range 20–50 meV. For either parameter, size or binding energy, no better agreement is expected of the effective-mass model.

3.4. Ionization energy and donor-level depth

Figure 7 illustrates the evolution of the μSR signals in CdS between 5 and 40 K (see footnote 6). Here the two satellite amplitudes have been summed as a measure of the Mu^0 yield and the amplitude of the central line taken as the Mu^+ yield. The overall sum remains essentially constant throughout, accounting for the full polarization of the incoming muons. One can say that Mu^0 ionizes or dissociates in the interval between about 20 and 30 K, converting to Mu^+. Alternatively, one can argue that thermal energy competes with Coulomb attraction to decide the probability of electron capture. This is still a moot point, as discussed in appendix C.

Analysis of the detailed temperature dependence requires a model for the defect statistics. Two common models are expressed in equations (6) and (7), where x is in each case the increment in signal amplitude, suitably normalized:

$$x \propto \exp(-E_i/kT); \tag{6}$$

$$x^2/(1-x) \propto T^{\frac{3}{2}} \exp(-E_D/kT). \tag{7}$$

A straight Arrhenius plot according to equation (6) gives an ionization energy of $E_i = 22 \pm 4$ meV, as shown in figure 8(a). This model would be appropriate to direct loss of electrons from the donor level to the conduction-band minimum, in which case E_i may itself be equated with the donor depth. If instead a Fermi level is considered to be established between the donor level and the conduction-band minimum, the defect level population should follow equation (7), which describes the fitted lines in figures 7 and 8(b). A derivation of this equation and discussion of its relevance to muonium statistics is given in appendix C. The resulting donor depth, $E_D = 49 \pm 5$ meV, is then rather more than twice the Arrhenius activation energy E_i. It is also about twice the values originally reported by Gil *et al* (1999,

[10] We use this term in preference to the other commonly used expression, namely *hydrogenic donor*. The latter seems somewhat of a misnomer in this context, since no such shallow state had previously been seen, either for muonium or for hydrogen itself!

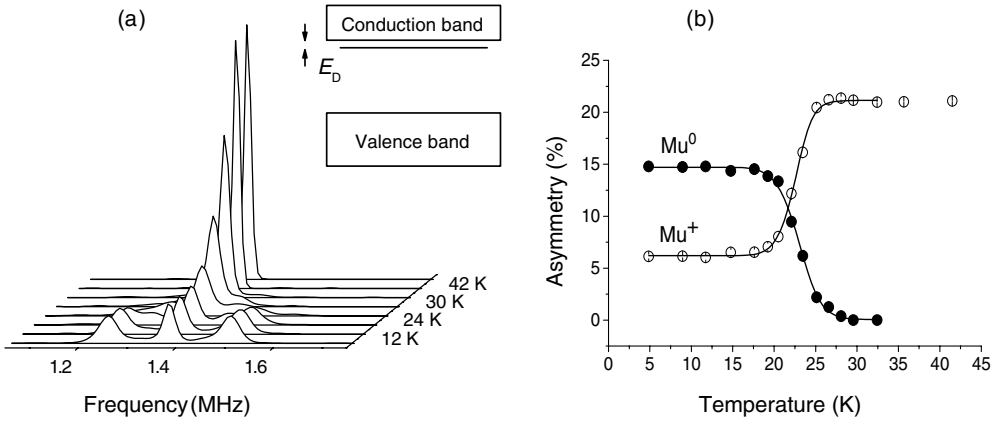

Figure 7. Evolution of the μSR spectra for CdS and time-domain analysis of the Mu^0 and Mu^+ signal amplitudes leading (via model fits of the temperature dependence, see figure 8) to the donor depth.

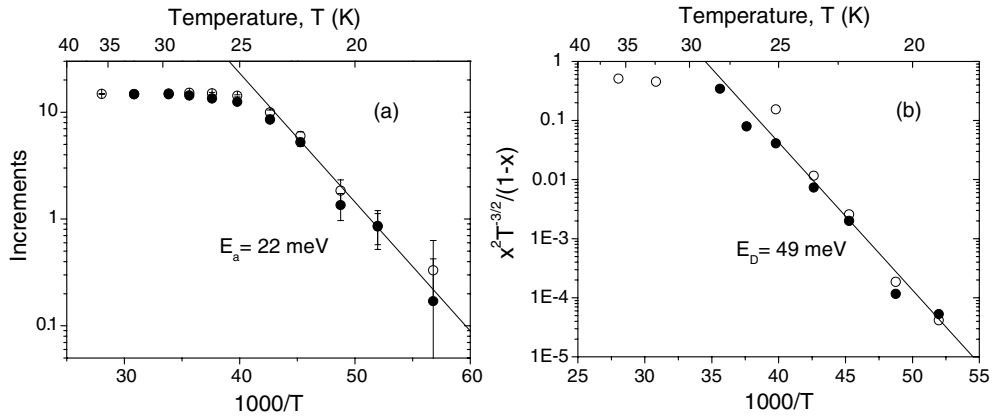

Figure 8. Alternative treatments of the amplitude data for CdS using (a) equation (6) and (b) equation (7). These correspond, respectively, to one-way ionization and dynamic (capture and loss) equilibrium. Variation of the paramagnetic and diamagnetic fractions (filled and open data points) give the same activation energy.

2001) for a different sample; this variation remains to be investigated but almost certainly represents interaction with native defects or impurities, at least in the earlier data. Pending resolution of these questions, the donor depths must be considered uncertain to within a factor of two or three. The various results are entered in table 5.

3.5. Local electronic structure

The microscopic structure envisaged for the shallow muonium centre is sketched in figure 9(a). Here the dilated 1s function is envisaged as an envelope function (see, for example, Yu and Cardona 1996) defining the superposition of atomic orbitals which carries the unpaired electron spin. These latter will have predominantly 5s(Cd) character in CdS, as appropriate to the conduction-band minimum. In other words, the electronic wavefunction is made up

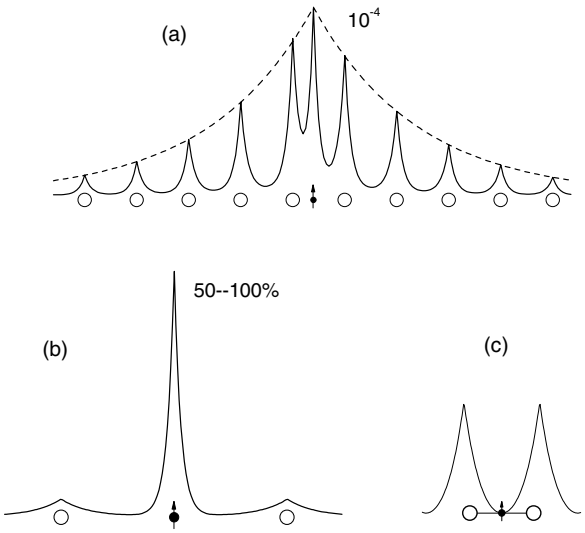

Figure 9. Local electronic structure of an effective-mass donor (a). This is envisaged superposition of 1s(Mu) and CB states comprising mainly cation s orbitals defined by an envelope function which resembles a dilated 1s wavefunction, with effective Bohr radius a^*. Central spin density is less than in a free muonium atom by a factor close to 10^4. Whether this is the maximum of the distribution (dashed curve) or not remains to be determined. This extended orbital is to be contrasted with the compact electron wavefunctions for the trapped-atom Mu_T^0 (b) and deep-donor Mu_{BC}^0 (c) states in Si and GaAs, etc (see appendix A).

of conduction-band states, in a package localized about the charge defect. Such an extended and composite orbital contrasts strongly with the compact electron wavefunctions of both the trapped atom (Mu_T^0) and the distinctive deep-donor (Mu_{BC}^0) states exemplified by muonium in silicon, sketched in figures 9(b) and (c).

In the following, we use the terms *deep* to designate centres with compact (defect-like) electronic wavefunctions and *shallow* for those with extended (bandlike) wavefunctions that approximate to the hydrogenic or effective-mass model. Binding energies or donor depths will usually be substantially less than 100 meV for shallow centres and substantially more so for deep centres. We note that some compact states can nonetheless be fully ionized at room temperature, and that this is indeed the case for Mu_{BC}^0 and H_{BC}^0 in silicon. It will also be useful to follow the usage of, for example, Yu and Cardona (1996), and retain the designation *deep* for compact centres that are band resonant, i.e. auto-ionizing.

It is certainly important to the shallow-donor interpretation to be sure that the low hyperfine parameters in CdS do not represent some compact molecular-radical type which happens to give low electron density at the muon site, e.g. in the manner of Mu_{BC}^0 in silicon (figure 9(c)) where the muon is located exactly at a node in the singly occupied molecular orbital (Cox and Symons 1986). One can also question whether spin density is maximum at the muon site as drawn in figure 9(a), i.e. whether it is described by the same envelope function which controls the spin density on surrounding nuclei. If not, the above estimates of the radius of the envelope function must be modified. The low value of the dipolar hyperfine parameter in CdS compared with Si is already reassuring in this respect (see appendix B) but the most compelling evidence for extended states comes in fact from measurements of the electron g-values, described in section 4. In principle the spin-density distribution could be mapped out by measuring superhyperfine interactions on those Cd atoms which carry nuclear spin,

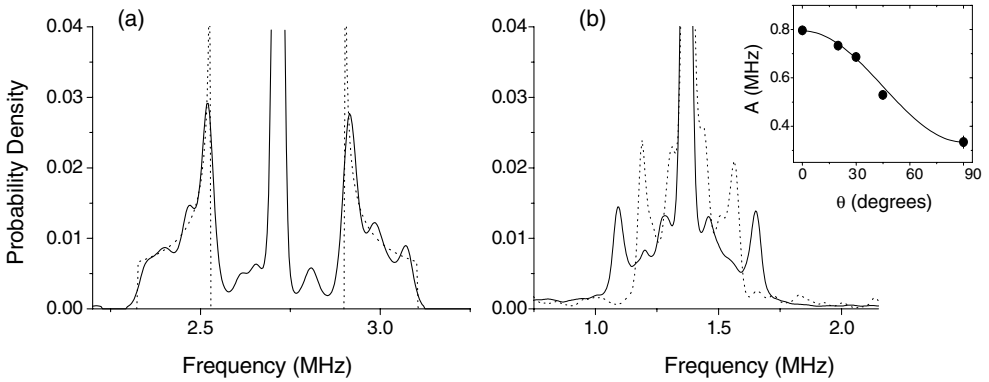

Figure 10. μSR spectra of the shallow-donor Mu^0 state in (a) polycrystalline and (b) single-crystal ZnO (from Cox *et al* 2001a, and Alberto *et al* 2001). The dotted curve in (a) is the powder-pattern lineshape generated by equation (29) (appendix B). The inset to (b) shows the orientation dependence of the main splitting, fitted with equation (28). Parallel independent work is reported by Shimomura *et al* (2002).

namely ^{111}Cd and ^{113}Cd, which together have a natural abundance of 25%. Such a mapping would provide a definitive picture of the electron wavefunction. In practice it has so far proved impossible to resolve individual interactions although decoupling data are broadly consistent with the expected distribution of values (Lord *et al* 2001).

3.6. Site and mobility

Since the atomic-like or trapped-atom Mu^0 state (figure 9(b)) is highly mobile in most semiconductors and insulators, even exhibiting an enhanced quantum diffusion at low temperature in materials such as GaAs and NaCl (Kadono 1990, Kadono *et al* 1990), one can ask whether a shallow-donor Mu^0 centre (figure 9(a)) could also be mobile. This seems unlikely, though not impossible. The trapped-atom state has the muon charge tightly screened by a normal 1s electron and consequently induces only a negligible elastic distortion of the host lattice. For the shallow-donor state, however, the screening is only effective at long range. Locally, i.e. on the scale of the elemental diffusion step, screening is minimal; the muon should then be immobile, as the Mu^+ centre usually is at these temperatures. For comparison, Mu^+ becomes mobile only above about 200 K in GaAs and other III–Vs with zincblende structure (Chow *et al* 1997). On the other hand, the unblocked channel sites within the wurtzite structure may facilitate Mu^+ or H^+ diffusion at lower temperatures. There is accumulating evidence of this for Mu^+ in III-nitrides (Lichti *et al* 2001, Lichti 2003a, 2003b) though it has yet to be considered explicitly for the II–VIs. The hyperfine anisotropy of the observed Mu^0 centre in CdS tends to suggest a static centre; certainly there is no spherical averaging. Pending an explicit site determination in the neutral state, it seems reasonable to suppose that this is the same as for the ionized donor of figure 5, i.e. adjacent to the sulfur anion in an AB or BC site. Somewhat surprisingly, sites with *c*-axis symmetry are found to be populated in slight preference to those in the off-axis sites.

4. The shallow muonium and protium states in ZnO

Figure 10(a) shows the μSR spectrum of polycrystalline ZnO, from which the shallow-donor state of muonium in this material was originally inferred. A distinctive powder pattern can be

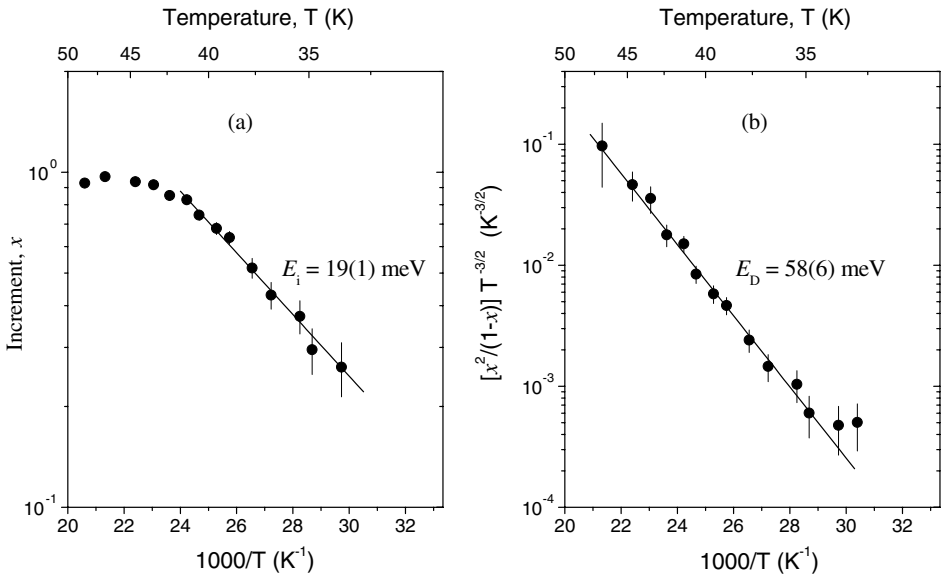

Figure 11. Ionization data in ZnO, analysed according to the alternative statistical models of (a) equation (6) and (b) equation (7) (from Cox *et al* 2001b).

discerned which gives principal values of the hyperfine tensor directly (an analytical form is given in appendix B). Such powder patterns are commonplace in ESR spectroscopy but this seems to be the first example in μSR—the low values of the parameters keeping the overall width down and making the Paschen–Back regime easily accessible.

4.1. Symmetry axis versus site determination

The single-crystal spectra obtained subsequently are shown for comparison in figure 10(b). As well as giving identical hyperfine parameters, their detailed angular dependence defines the symmetry axis of the centre as being parallel to the crystallographic *c* axis. This result is to be compared with the possible proton sites adjacent to the oxygen anion, sketched in figure 4. Van de Walle (2000) found all four to have very similar energy, with $BC_\perp$ favoured by a small margin; this prediction of the most stable site has since been revised to $BC_\parallel$, i.e. the BC site with *c*-axis symmetry (Lavrov *et al* 2002). By implication, all four sites could accommodate shallow H^0 or Mu^0 centres. The μSR data in fact indicate sites with *c*-axis symmetry to be even more preferentially populated in ZnO than in CdS: Alberto *et al* (2001) found no need to include occupation of off-axis sites in ZnO from their analysis of the spectra of figure 10(b). Their data do not distinguish between the BC and AB *c*-axis sites, however, i.e. between $AB_{O,\parallel}$ and $BC_\parallel$. Shimomura *et al* (2002) claim identification of two distinct shallow Mu^0 centres in ZnO, one in each of these sites. In the absence of information on the ^{67}Zn nuclear couplings, however, it is hard to see how these assignments can be made so categorically.

4.2. Activation energies versus donor-level depth

An ionization region in which the Mu^+ central line grows at the expense of the Mu^0 satellites is apparent in ZnO, as in CdS. Figure 11 shows the two alternative methods of analysis. The straight Arrhenius plot of figure 11(a) and equation (6) gives an ionization energy of 19 ± 1 meV,

improving on the value of 30 ± 5 meV from the sparser original data on a sample from the same batch (Cox *et al* 2001a). The assumption of thermodynamic equilibrium in the analysis of figure 11(b) and equation (7) gives instead a donor depth of 58 ± 6 meV. The approximate factor of two between the different methods comes from the asymptote of $x^2/(1-x)$ in equation (7), somewhat enhanced by the $T^{3/2}$ density-of-states term. Shimomura *et al* (2002) measure a direct activation energy of 25 meV, assigned to the $BC_{\parallel}$ site, and double this to quote a donor depth of 50 meV. As for the case of CdS, and despite the apparent precision of particular fits to particular data sets, it would be premature to claim that the donor level is known to better than a factor of three until the correct statistical model is established.

A second activation energy of 3 meV, reported by Shimomura *et al* (2002) and assigned to the $AB_{O,\parallel}$ site, is surely too low for the effective-mass model. For ZnO, with $\varepsilon \approx 8$ and $m^*/m = 0.24$ (Madelung 1996), the prediction of equation (5) is $R^* \approx 52$ meV. One can ask whether it might correspond to an electronic excited state, but the hyperfine parameters should then be much lower: in the hydrogenic series, binding energy varies as $1/n^2$ and contact interaction at the nucleus as $1/n^3$. 3 meV is also too low, i.e. too similar to kT, for Arrhenius analysis to be strictly valid. It corresponds to the weaker temperature dependence noted both in μSR data by Cox *et al* (2001a) and in ESR data by Hofmann *et al* (2002) and attributed in both cases to defect or impurity interactions. A similarly weak temperature dependence of the μSR amplitudes in the original CdS data was noted to be sample dependent by Gil *et al* (1999) but is absent in the newer data of figure 7.

4.3. Determination of electron g-values for hydrogen and muonium in ZnO

Figure 12 shows the conventional magnetic resonance spectra of hydrogen itself in ZnO, from the study due to Hofmann *et al* (2002). The proton-ENDOR spectrum correlates with one of two native-donor resonances in the ESR spectrum, confirming the role of hydrogen as a shallow donor in this material. The reported ionization energy of 35 ± 5 eV lies within the range bracketed by the μSR data of figure 11 but there is a puzzling discrepancy between the muon and proton hyperfine splittings, which we discuss in appendix B. Although this raises the question of whether the muonium and protium centres are entirely analogous, in this section we see that they have identical electronic g-values.

Figure 13 illustrates recent double-resonance measurements of the electronic g-values for the muonium donors, both in ZnO and CdS. Fast transitions of the electron spin are induced by radiofrequency magnetic fields at the ESR frequencies and the effect detected by a collapse of the hyperfine splitting in the μSR spectrum. The transitions involved are sketched in the figure. This novel technique differs from earlier muonium double-resonance experiments (Estle and Vanderwater 1983, Patterson 1988) by exploiting the small value of the hyperfine splitting to drive both ESR transitions simultaneously, effecting a dynamical decoupling of the muon and electron spins. This same dynamical decoupling can also give rise to a remarkable resonant *repolarization* of the muons, visible in low longitudinal fields (Lord 2003).

The result for muonium in ZnO is $g = 1.96 \pm 0.01$, less precise than the ESR measurement of figure 12(a) but entirely consistent with it. The point is that this is distinct from the free-electron value of 2.002 and equal to that of other chemical impurities which form shallow-donor states in this material (Block *et al* 1982). In fact, this value is believed to correspond to conduction-band electrons at the band minimum and so to be independent, within the effective-mass model, of the chemical nature of the charge defect. The equivalent resonance for CdS is broader, presumably because dipolar nuclei are 25% abundant in CdS, as against 5% in ZnO, but the g-value of 1.82 ± 0.06 is for this material even further from the free-electron value. Again, it corresponds closely to the value 1.78 known for other shallow donors in

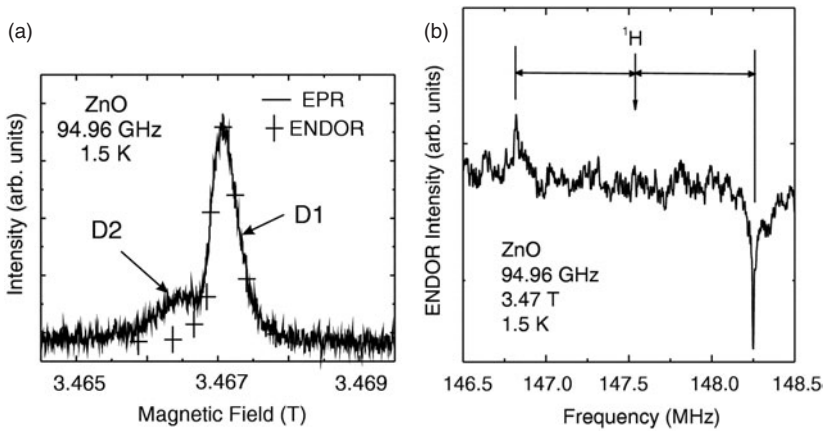

Figure 12. ESR and proton-ENDOR spectra for ZnO (Hofmann *et al* 2002). Two native paramagnetic centres were detected by these authors, overlapping in the usual X-band ESR spectrum at 9.5 GHz but well resolved at 95 GHz (a). The resonance labelled D1 could be convincingly assigned to a hydrogen shallow donor via its correlation with the proton-ENDOR spectrum (b). (Reproduced with permission from Hofmann *et al* (2002), copyright (2002) American Physical Society.)

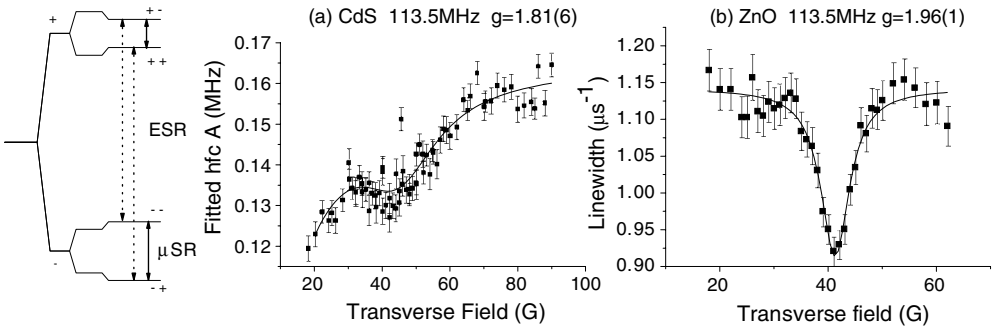

Figure 13. ESR–μSR double resonance. The energy level diagram is the same as that sketched in figure 6, showing here both the low-frequency ENDOR-type or μSR transitions and the high-frequency ESR-type transitions excited in the double-resonance measurement of electronic g-values. The electron spin resonance is detected via a reduction of the satellite splitting or rms width of the muon spin rotation spectra (Cox *et al* 2002a).

CdS (Patel *et al* 1981). Trigger detection of the ESR resonance via the μSR response may be compared in some respects to the optically detected magnetic resonance which gives the necessary sensitivity for measurements on the conventional shallow donors.

In the absence of a direct mapping of Mu^0 spin density via the individual ^{67}Zn, ^{111}Cd or ^{113}Cd couplings, these results may be taken as reasonable evidence of an extended electron orbital with conduction-band character. A similar determination for the shallow Mu^0 state in CdTe (see following section) would provide an especially stringent test, since other shallow donors have a *negative* g-value in that material (Simmonds *et al* 1982). Although the sign is not normally accessible, crossed-coil techniques, driving the ESR transition with a defined sense of precession, may do the trick.

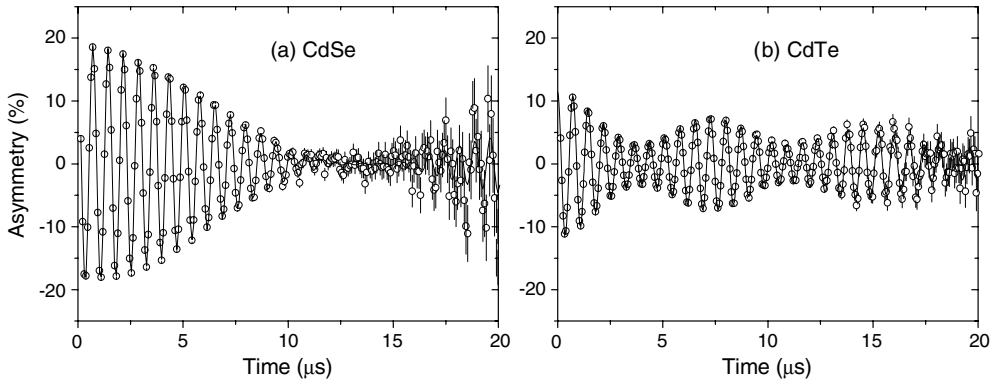

Figure 14. μSR precession signals for CdSe (a) and CdTe (b). Note the slow beating in (a), indicative of a very low hyperfine constant, and the missing polarization in (b), the result of a short-lived atomic-muonium state.

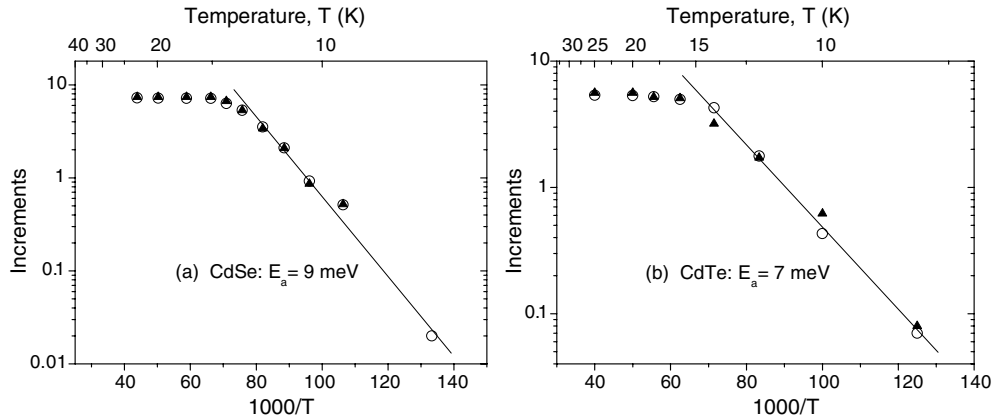

Figure 15. Arrhenius analysis of the ionization regimes for CdSe and CdTe, according to equation (6). As in figure 8, variations of paramagnetic and diamagnetic fractions give the same results (open and filled data points). Using the same data but with the equilibrium model of equation (7), Gil *et al* (2001) obtain donor depths of 20 and 16 meV, respectively.

5. II–VI shallow-donor survey

The readily available members of the II–VI family of compounds have by now all been examined by μSR spectroscopy. Mu^0 states with shallow-donor characteristics, similar to those described above in CdS and ZnO, have also been identified with certainty in CdSe and CdTe. Figure 14 shows their precession signals and figure 15 their ionization behaviour. The results for these four compounds are summarized in tables 4 and 5. Of the others, tantalizing hints of similar signals have been seen in powder samples of CdO and HgS but remain to be confirmed: the satellite amplitudes are so weak compared with the central Mu^+ line in these two candidates that it has so far proved impossible to extract reliable temperature dependences (Cox *et al* 2002a). No signs of shallow centres have been found in II–VI compounds where the cation is from group IIA, i.e. Be, Mg, Ca, Sr or Ba.

The raw precession signals of figure 14 both deserve comment. In CdSe, the hyperfine coupling is so small that the beats are only visible towards an elapsed time of 20 μs. This

Table 4. Hyperfine parameters of the shallow-donor muonium states (from Gil *et al* 2001, Davis *et al* 2003). Further data for ZnO are given in table B.2.

	A_{iso} (kHz)	A_{iso}/A_0	D (kHz)	a^* (exp) (nm)	a^* (theo) (nm)
CdS	244 ± 5	0.55×10^{-4}	91 ± 6	1.4	2.4
CdSe	87 ± 4	0.19×10^{-4}	<40	2.0	3.9
CdTe	261 ± 4	0.58×10^{-4}	<50	1.4	6.2
ZnO	500 ± 20	1.12×10^{-4}	260 ± 20	1.1	1.7
InN	92 ± 5	0.20×10^{-4}		1.9	4.0

illustrates the quality of data from the ISIS pulsed source, where the evolution of muon polarization can be followed as far as ten muon lifetimes. CdTe has the cubic (zincblende) structure, so the distinctive beating in this signal pre-empts any idea that the new shallow-donor muonium states are only to be found in hexagonal (wurtzitic) hosts. The initial amplitude of the signal is seen to be lower, however, so that not all the incoming muon polarization is accounted for in this transverse-field signal. The manner in which this missing polarization is recovered in longitudinal-field experiments (see appendix B) is indicative of normal muonium, i.e. of a trapped-atom state. It therefore appears that deep and shallow muonium states are *both* formed in CdTe. This is so far the only material to show this coexistence but it gives an important clue to the nature of the shallow-to-deep instability, expressed in figure 24. ZnS and ZnSe, on the other hand, show clear signals of the trapped-atom state, but no sign of the shallow donors. A different type of centre again is seen in HgO: this is presented in section 6, before the overall systematics are reviewed in section 6.

6. A deep-donor muonium state in HgO[Note 11]

μSR signals and spectra in ZnO and HgO are shown side by side in figure 16. Superficially they look similar, until the time and frequency axes are inspected closely. Then it is seen that both the contact and dipolar hyperfine parameters are some 20–30 times greater in HgO. The two methods of analysing the temperature dependence of satellite amplitudes likewise give ionization energies and temperatures an order of magnitude higher than those in ZnO, as shown in figure 17. All these parameters are far too large for the effective-mass model for which, with $\varepsilon = 10$ and $m^*/m = 0.09$ (Madelung 1996) $R^* = 12$ meV. They resemble much more closely those of the Mu^0_{BC} centres known in the group-IV elements Si, Ge and diamond and the III–V compounds GaAs and GaP (table 5).

In appendix A, Mu^0_{BC} in Si is identified as a deep donor: it has a compact electron wavefunction, sketched in figure 9(c), which already borrows AB character from the host (Cox and Symons 1986) and its direct ionization, without site change, defines a donor depth some 210 meV below the CB edge (Hitti 1999). Whether the muon site is bond centred in HgO remains to be established. The crystallographic structure comprises zig-zag chains which also allow for sites that are AB to the oxygen anions, as well as bridging sites between chains. Nonetheless, it seems reasonable to assign donor characteristics to this Mu^0 centre, contrasting it with a possible acceptor function of the trapped-atom state in other oxides. Its characteristics are favourable to another form of spectroscopic detection, namely level crossing resonance in a longitudinal magnetic field, illustrated in appendix B.

[11] At the time of writing, there are preliminary but clear indications of a similar state in Ag_2O. Details will be published elsewhere.

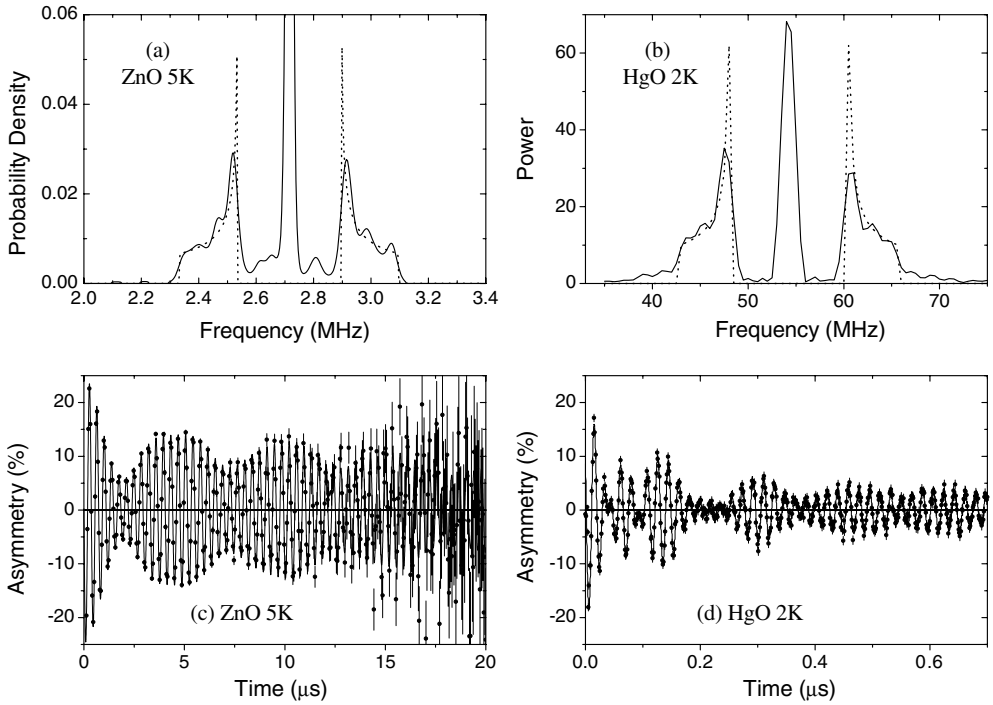

Figure 16. μSR frequency spectra and corresponding time-domain signals in ZnO and HgO (from Cox *et al* 2001b). Note the different frequency and timescales! (Reproduced with permission from Hofmann *et al* (2002), copyright (2002) American Physical Society.)

Table 5. Ionization parameters. The two alternative values for the donor depth, E_i and E_D, are determined from temperature-dependent data, using equations (6) and (7), respectively. The binding energy R^* (experimental) is inferred from the low-temperature hyperfine constants using equation (32) (appendix B). The expectation of the effective-mass model, R^* (theoretical), is given by equation (5). The first two entries for CdS represent different data sets on the same sample; the third is for a newer sample. The two entries for CdSe are from the same data set, as are those for CdTe. The first two entries for ZnO are from different data sets on powder samples from the same batch. The second entry for Shimomura *et al* refers to an additional assignment.

	Reference	$T_{ionization}$ (K)	E_i (meV)	E_D (meV)	R^* (exp) (meV)	ε (m_e)	m^* (meV)	R^* (theo)
CdS	Gil *et al* (1999)	22	9	18		8.8	0.22	33
CdS	Gil *et al* (2001)	22		26 ± 6	59	8.8	0.22	33
CdS	Figures 7, 8	22	22 ± 4	49 ± 5		8.8	0.22	33
CdSe	Gil *et al* (2001)	12		20 ± 4	37	9.7	0.12	19
CdSe	Figure 15	12	9 ± 2			9.7	0.12	19
CdTe	Gil *et al* (2001)	13		16 ± 4	52	10.2	0.09	11
CdTe	Figure 15	13	7 ± 2			10.2	0.09	11
ZnO	Cox *et al* (2001a)	38	30 ± 5		79	8.3	0.26	52
ZnO	Figure 11	38	19 ± 1	58 ± 6	79	8.3	0.26	52
ZnO	Shimomura *et al* (2002)	28	25	50		8.3	0.26	52
ZnO(2)	Shimomura *et al* (2002)	18	3	6		8.3	0.26	52
InN	Davis *et al* (2003)	40	12 ± 1		30	9.1	0.12	20

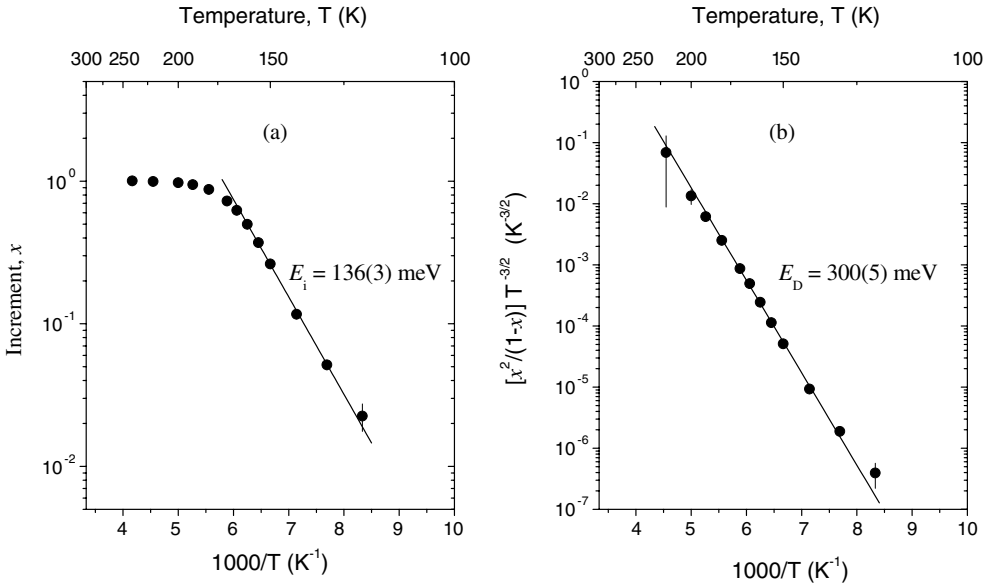

Figure 17. Mu^0 ionization in HgO, analysed in the two models of (a) equation (6) and (b) equation (7).

7. Nature of the shallow-to-deep instability

We consider now how it is that, even within the same II–VI family of materials, muonium can show examples of shallow-donor, deep-donor and trapped-atom configurations. We keep in mind which types of centre exhibit coexistence in a given material and which appear to be mutually exclusive, seeking a universal description which will carry over, *mutatis mutandi*, to hydrogen in a wide variety of semiconductors.

7.1. Host properties versus impurity properties

Whereas it is never simple to predict the behaviour of any particular impurity, interstitial hydrogen and muonium appear to satisfy all the usual criteria expected to favour deep-level centres with compact electronic orbitals. These are (see, for instance, Stoneham 1975, Yu and Cardona 1996) the presence of s electrons (as in Mu_T^0 and the other trapped Mu^0 atoms), or of large local distortions of the lattice (as for Mu_{BC}^0 in Si), or of a highly localized potential (as the unscreened Coulomb potential of the proton or muon indeed is, there being no core electrons resembling those of the host atoms).

The shallow-donor muonium states seem to defy these guidelines! In order to construct a localized orbital from CB states, deep or compact states require the full spread of wavevectors. This is more costly in energy when the CB is wide than when it is narrow, so that one can imagine deep states to be precluded in wide-band materials. By the same token, the low effective mass of electrons at the minimum of a wide band favours shallow states: the more extended shallow states need only a small spread of wavevectors around the CBM. Yet shallow muonium states are found in CdS and ZnO, both of which are quite ionic, with narrow bands. Not even the huge lattice distortion in ZnO, estimated by Van de Walle (2000) as a bond-stretch of some 40% for H_{BC}^+, is able to constrain the electron in a compact orbital. Another strange contradiction is GaAs, which has a large dielectric constant, small electron effective mass

and *very small* Fröhlich coupling constant (Stoneham 1975), leading to extremely shallow substitutional donors; yet the interstitial muonium states, both cage centred and bond centred, are deep and compact. For the specific case of H and Mu in Si and Ge, Altarelli and Hsu (1979) attribute the breakdown of effective-mass theory and the formation of deep-level compact states to large valley–orbit interactions (contrasting with the shallow state formed by interstitial Li). More generally, Herbert and Inkson (1977) find these interactions to control shallow-to-deep transitions, even exhibiting a threshold behaviour. Remarkably similar to the muonium case, is that of halogen impurities in II–VIs: substituting for the anion, F gives a deep-level or compact centre in BeO and ZnS whereas I, Br and Cl all give shallow donors in CdS (Stoneham 1975). A clue to the similarity (Stoneham 2001) may be that protonation of the anion effectively changes the anion charge state by one unit, as does its substitution by F^-, Cl^-, etc:

$$O^{2-} + H^+ \rightarrow OH^-. \tag{8}$$

Structure and coordination may be important here. In a covalent oxide such as silica, the twofold coordinated oxygen can use its lone electron pairs to form a dative bond to protons, weakening but not breaking the silicon–oxygen bonds. In the II–VIs, the anions are fourfold coordinated and have no lone pairs. Protonation of the anion then implies breaking a bond to a cation, and a greater local distortion in consequence. We return to these chemical aspects in sections 9 and 10, after a further examination of the systematics.

7.2. Role of electron affinity

Of all the potentially relevant material properties, none of those which are precisely known or commonly tabulated—e.g. bandgap, ionicity, dielectric constant, bond-length, crystal structure etc—shows a simple threshold for the switch from deep to shallow behaviour. Although reliable values are hard to come by, there is increasing evidence that it is not so much the overall width of the conduction-band but the depth of its minimum below the vacuum continuum, i.e. the electron affinity, that is the elusive controlling parameter.

Figure 18(a) shows a band-offset diagram expressing the importance of electron affinity. This qualitative sketch is reproduced from Davis *et al* (2003) and is a relevant subset of more extensive band-offset diagrams shown previously in conference presentations by Van de Walle (2000, 2001). Fuller quantitative versions are now available (Van de Walle 2003, Van de Walle and Neugebauer 2003). The dotted line is envisaged as the H(+/−) transition level of figure 3, i.e. the value of the Fermi energy at which the positive and negative charge states of hydrogen impurity are in equilibrium. This is the so-called pinning level, i.e. the level at which the Fermi energy would be pinned when hydrogen defects were in abundance. (Negative U is tacitly assumed.) When the pinning level lies deep in the gap, amphoteric compensating behaviour results. When it intersects the conduction-band, any occupied compact state is band resonant and therefore auto-ionizing. The electron cannot be bound in a compact orbital but falls to the CBM; it is available for charge transport at ordinary temperatures or is held weakly in an extended shallow state at cryogenic temperatures. This latter is the neutral or undissociated state of the effective mass donor. The same concept is expressed by Kiliç and Zunger (2002), working with a subset of oxides, and similar pinning rules for transition-metal impurites in III–V and II–VI semiconductors had been noted previously by Zunger (1986). These various authors use density-functional calculations and find that, for those materials in which they have calculated the H(+/−) level explicitly, it lies at closely similar—though not identical—depths below the vacuum level. They extrapolate these findings to suggest that the level may be essentially common to a wide variety of materials and use it to predict hydrogen-induced conductivity wherever it intersects the conduction-band.

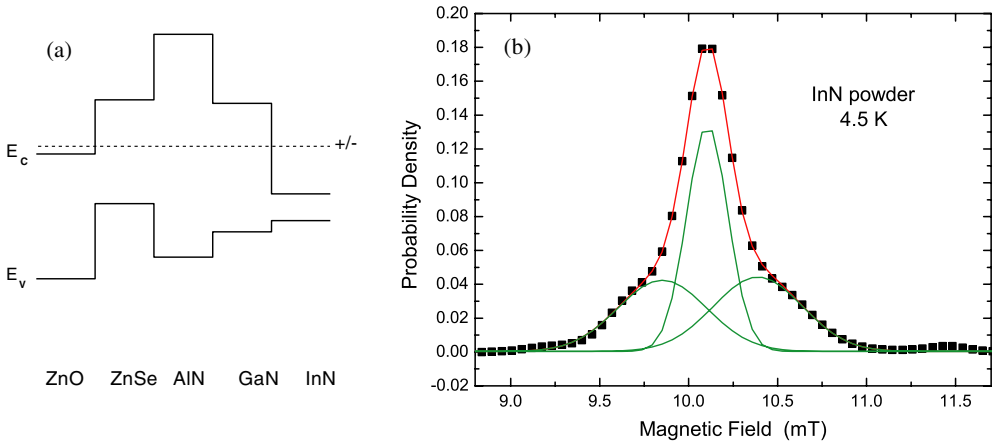

Figure 18. Band offset diagram (a) and μSR spectrum (b), respectively predicting and revealing a shallow-donor Mu^0 state in InN (reproduced with permission from Davis *et al* 2003: copyright (2003) American Physical Society).

(This figure is in colour only in the electronic version)

7.3. InN: a second successful prediction

Referring to figure 18(a), the contrast between the deep level in ZnSe and the band-resonant level in ZnO emerges explicitly in the calculations described by Van de Walle (2002). The compact trapped-atom Mu^0 state in ZnSe was already known and the extended shallow-donor state in ZnO discovered in direct response to this prediction. A similar contrast between the III–V materials GaN and InN is described by Limpijumnong and Van de Walle (2001). For GaN, the predictions of deep-level amphoteric behaviour and strongly negative U seem also to be borne out by μSR studies: Mu^+ and Mu^- spectra are readily assigned but there are only indirect indications of a short-lived Mu^0 state (Cox *et al* 2002b, plus references therein).

The prediction of shallow-donor behaviour for InN implied in figure 18(a) prompted another successful μSR search. Figure 18(b) depicts the resultant spectrum for polycrystalline InN (single-crystal samples of this material are not yet available). In this III–V compound, all naturally abundant nuclei carry sizeable magnetic dipole and electric quadrupole moments. This gives rise to a damping of the muon spin rotation signals and a corresponding broadening of their frequency transforms, already evident in the signal for the electronically diamagnetic Mu^+ state. Even allowing for quadrupole splittings, the unusual lineshape of the low-temperature spectrum of figure 18(b) cannot be explained entirely in these terms, however; it evidently subsumes an unresolved hyperfine splitting corresponding to a Mu^0 state. The three components of the signal may be separated via time-domain fitting, the central and satellite amplitudes varying with temperature as in figure 19(a) and defining an ionization regime centred around 40 K. The Arrhenius plot of figure 19(b) gives a direct ionization energy for this centre of 12 meV.

It is apparent, then, that bandgap does not play a controlling role, this being much lower in InN than in the II–VI compounds where the other shallow states have been found. In fact, the bandgap in InN has recently been revised downwards, from the previously accepted value of 2 eV to about 0.8 eV (Wu *et al* 2002). InN provides the first example of a shallow muonium state in a semiconductor which is not classed as a wide-gap material, i.e. which is not transparent to some or all wavelengths of visible light. After ZnO, it is the second example

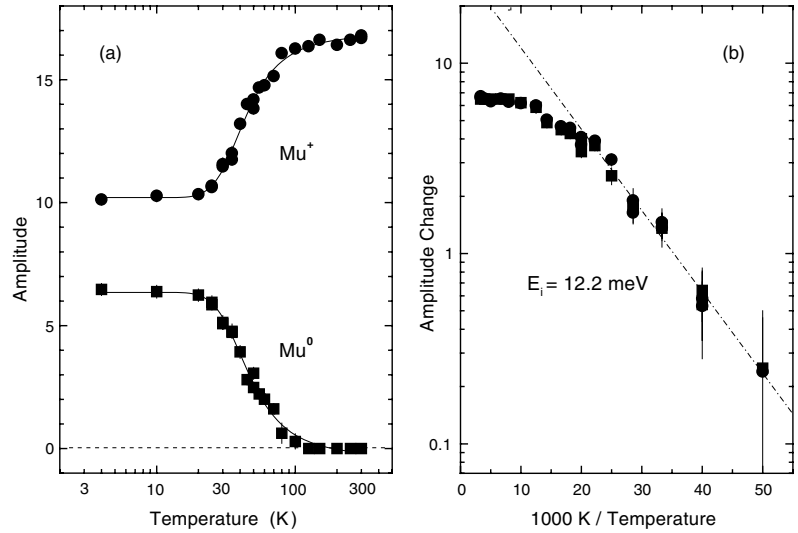

Figure 19. Temperature-dependent amplitudes of the Mu^0 and Mu^+ signal components in InN and an Arrhenius plot (equation (6)) of changes through the ionization region (reproduced with permission from Davis *et al* 2003: copyright (2003) American Physical Society).

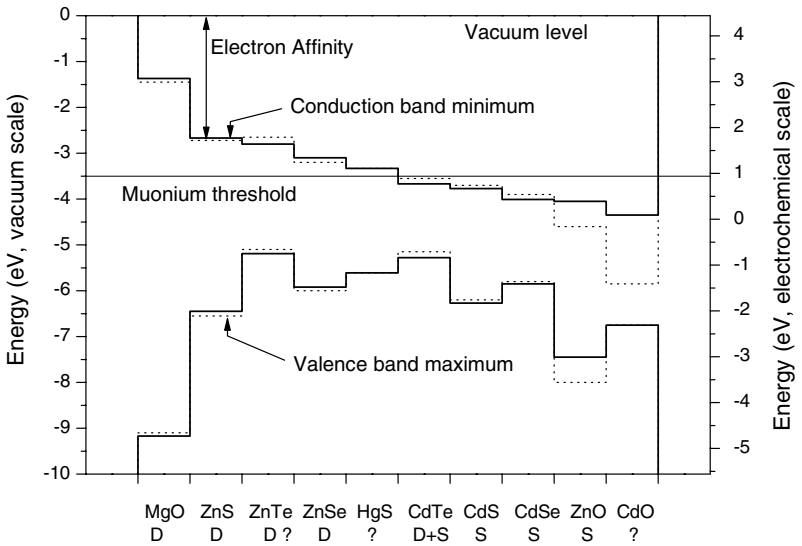

Figure 20. Band-offset diagram for the II–VIs, indicating limits for the threshold electron affinity, apparently controlling the switch from deep (D) to shallow (S) muonium behaviour. Question marks indicate results to be confirmed; the others are well established.

of muonium studies confirming a theoretical prediction for hydrogen. As in the case of ZnO, it would also explain a longstanding puzzle of native n-type conductivity in InN and is entirely consistent with recent electrical measurements (Look *et al* 2002).

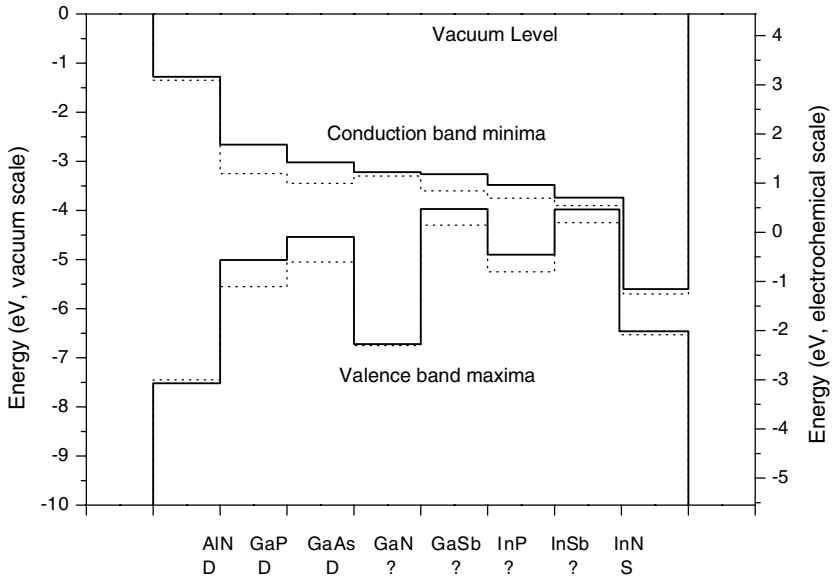

Figure 21. Band-offset diagram for the III–Vs. Candidate pinning levels all intercept the CB in InN, leading to shallow-donor behaviour; that of Van de Walle and Neugebauer (2003) also intercepts the VBs of GaSb and InSB, suggesting shallow-acceptor behaviour.

8. Band-offset diagrams for II–VIs and III–Vs

8.1. Threshold behaviour in the II–VIs

Figures 20 and 21 show how the currently available μSR data, for the II–VIs and III–Vs respectively, can be taken as supportive of an electron-affinity threshold. The solid histogram of figure 20 is the band-offset diagram shown by Cox *et al* (2002a), constructed as follows: relative values of the valence-band (VB) maxima are taken from the comprehensive calculations of Wei and Zunger (1998) and aligned on the vacuum scale using Kiliç and Zunger's (2002) values for ZnO and MgO. Empirical values of the bandgaps, as tabulated by Madelung (1996)[12], are then added to give the CB minima. Their distances below the vacuum level are a measure of electron affinity for each material—an elusive parameter for which individual absolute values are in principle indeterminate. We are concerned here with a bulk property of each semiconductor so that, although mixing theoretical and experimental data, this procedure at least avoids the uncertainties which plague electrochemical measurements of electron affinities. (These latter are made on heterojunctions or exposed surfaces and are subject to numerous artefacts and corrections; see e.g. Yu *et al* (1992), Fall *et al* (2002).) The dotted histograms are the recently published band lineups due to Van de Walle and Neugebauer (2003), superimposed without adjustment.

With increasing depth of the CB edge, a quite precise threshold appears to control the switch from deep (D) to shallow (S) behaviour for muonium. A lower limit of -3.6 eV is consistent with the established shallow behaviour in CdTe, CdS, CdSe and ZnO. It would also be consistent with a shallow state in CdO, if this is confirmed. An upper limit of -3.2 eV must be taken if a shallow state is also confirmed in HgS. The electron affinities themselves

[12] Except (a) InN, where the new value due to Wu *et al* (2002) has been used, and (b) ZnO, where Kiliç and Zunger appear to use the direct gap rather than the smaller indirect gap.

are certainly not reliable to this precision, so the muonium threshold might be quoted as, say, -3.5 ± 1 eV.

8.2. Shallow acceptors in the III–Vs?

Whilst InN is so far the only clear example of a shallow muonium donor amongst the III–Vs, figure 21 shows how a similar threshold could apply in this family. This diagram is constructed in the same manner as figure 21, with the relative VB edges of Wei and Zunger (1998) in this case aligned on values for the three nitrides given by Van de Walle and Neugebauer (2003). Deep-donor Mu^0_{BC} states are well known in GaAs and GaP, where they coexist with the trapped-atom Mu^0_T states that may be associated with a deep-acceptor function—the corresponding negative ion, Mu_T, is well characterized in n-type GaAs, for instance Chow *et al* (1995). An atomic-like Mu^0 state is also evident in AlN (Cox *et al* 2000). In the other materials of figure 21, charged muonium states predominate, as exemplified by studies in GaN (Cox *et al* 2002b, Lichti 2003a) and InP (Lichti *et al* 1997).

All of the three candidate thresholds—the -3 eV pinning level of Kiliç and Zunger, the -3.5 eV empirical muonium threshold of figure 20 and the -4.5 eV pinning level of Van de Walle and Neugebauer—are consistent with the known muonium behaviour. Whereas the first two lie above all the valence-band maxima, however, intersection of the latter with the valence-bands of GaSb and InSb raises the question of whether H and Mu might form *shallow-acceptor* states in these two materials. It seems unlikely that spectroscopic signatures of the neutral acceptors will be detected as easily as those of shallow donors, however, owing to the greater spin–orbit and spin–lattice interactions at the valence-band maxima. Their μSR signals would simply be depolarized and absent below the ionization temperature[13]. Magnetic resonance or μSR spectroscopy may in these cases be no substitute for electrical measurements.

9. An ionic model for the pinning level

In view of the overall consistency with the muonium data, and especially the success of the ZnO and InN predictions, it is tempting to identify the theshold electron affinity in figures 20 and 21 with the H(+/−) pinning levels of Van de Walle (2001, 2003) or Kiliç and Zunger (2002). In this section we examine how the concept of a common level might be understood, and how it is related to the well known, but very different, ionization energies for hydrogen in the vacuum state,

$$I(0/+) = 13.6 \text{ eV} \tag{9}$$

and

$$I(-/0) = 0.8 \text{ eV}. \tag{10}$$

These are the Rydberg, i.e. the binding energy of a single electron in the free hydrogen atom, and the binding energy of the second electron in the free hydride ion, respectively. Just how much these energies are modified when hydrogen is embedded in a solid medium is apparent from the fact that, if it becomes a negative-U system, the second electron is bound to the neutral defect more strongly than is the first electron to the interstitial proton! This can be the case even in a non-polar lattice such as silicon, for which the possibility of negative-U behaviour is explained in appendix A. In the following, we assume a purely ionic lattice (e.g. with nominal ionic charges 2+ for Zn and 2− for O in ZnO) and adapt procedures due to Stoneham and

[13] In InSb, the bandgap is so small and the spin–orbit coupling so large that any shallow-state hyperfine signature would be impossible to resolve in the presence of nuclear dipolar interactions, be it in a donor or acceptor state.

Ramos (1993) from the case of a substitutional impurity to that of an interstitial. The principal difference here is the much greater choice of crystallographic position and consequent variation of Madelung energy (Cox 2003a). The transition points of figures 2 or 3 may be written as

$$E(0/+) = W_+ - W_0 - I(0/+) \tag{11}$$

and

$$E(-/0) = W_0 - W_- - I(-/0), \tag{12}$$

with

$$W_+ = W_0 + (+1)A_+ - (+1)^2 B_+ = W_0 + A_+ - B_+ \tag{13}$$

and

$$W_- = W_0 + (-1)A_- - (-1)^2 B_- = W_0 - A_- - B_-. \tag{14}$$

Here the $W_\pm$ are, so to speak, ion insertion energies (in place of the ion replacement energies of Stoneham and Ramos), W_0 is an elastic distortion energy and the $A_\pm$ and $B_\pm$ are Madelung and polarization energies, respectively, these latter being proportional to the square of the charge defect. Using equations (2), (9) and (10), the pinning level is given by

$$E(-/+) = \tfrac{1}{2}(A_+ + A_-) - \tfrac{1}{2}(B_+ - B_-) - 7.2\ \text{eV}. \tag{15}$$

There are too many uncertainties to proceed to a numerical estimate but the point here is that the pinning level appears in this model as an opposition of quite large positive and negative terms. An upper estimate for the quantity A_+ in ZnO, for instance, is given by the energy cost of changing the anion charge state by one unit (as in equation (8)):

$$\frac{2e\alpha}{4\pi\varepsilon_0 r} \approx 24\ \text{eV} \tag{16}$$

(the Madelung constant is $\alpha = 1.64$ for wurtzite structures; the bond-length is $r \approx 200$ pm for ZnO). The quantity B_+ is essentially the OH bond-strength, since the OH bond is broken on going from H^+ to H^0 or H^-. The homolytic bond-strength is generally given as 470 kJ mol^{-1}, i.e. 4.9 eV, and serves as a lower limit: the heterolytic bond-strength is greater still. Both H^+ and H^- will also induce a lattice polarization, via ionic displacements, which may be estimated by a cavity model to be of order 1/10 Hartree or 2.7 eV. Being proportional to the square of the charge defect, this cancels in equation (15) (although it contributes to the Hubbard or Anderson U, equation (3)). Both the Madelung energies and bond-strengths decrease as the lattice parameter and anion size increase, in the series ZnO, ZnS, ZnSe, ZnTe, for instance, giving a surprising degree of compensation in equation (15). Treating the proton as a separate ion, in the loosely defined AB site adjacent to the anion, leads to a Madelung term A_+ with a fierce dependence on the precise position. There will likewise be some degree of freedom in A_-, although the bulky hydride ion will presumably be constrained to a more symmetric and so less sensitive position. Quite small adjustments to the H^+ and H^- sites clearly provide a means of fine-tuning the pinning level to a value common to these and perhaps other materials.

Nonetheless, the concept of an essentially universal value, common to a wide range of materials, seems surprising. Anticipating extrapolation to other oxides, the situation may be different for heavier cations, such as Cu^+ or Ag^+, which have a significant proton affinity (see, for example, Symons 1995). The cation may then compete with the anion for protonation, inverting the above balance of Madelung energies.

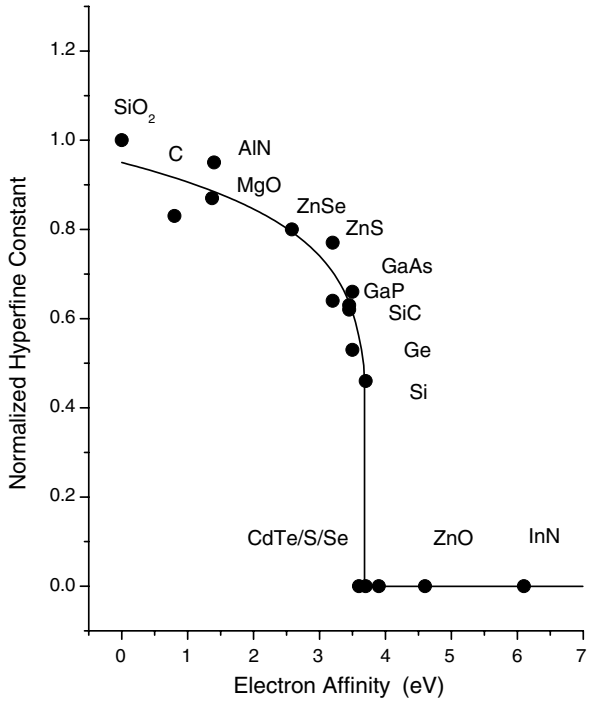

Figure 22. Correlation of muonium hyperfine constant (normalized to that of the free atom: values for the shallow-donor states are $\leqslant 10^{-4}$) with electron affinity, i.e. depth of the CBM.

10. Molecular orbital considerations

It remains to find some more intuitive understanding of the shallow-to-deep transition and the further predictive capability of an electron affinity threshold. Figure 22 shows just how sharp this transition is. Here the muonium hyperfine constant, taken as a measure of 1s(Mu) content or spin density on the muon, is correlated with electron affinity. Values for the trapped-atom states, to the left of the apparent transition, vary from essentially 100% of the free-atom or unperturbed value in quartz to about 50% in silicon. Those to the right of the transition correspond to the shallow-donor states where the values are four orders of magnitude smaller (table 4). The electron affinities, i.e. the depths of the conduction-band minima below the vacuum level, are taken from the band-offset diagram due to Van de Walle and Neugebauer (2003), so as to obtain a consistent set of relative values[14].

Figure 22 suggests increasing dilation of the electron wavefunction as the conduction-band edge is lowered until, at a critical electron affinity close to 4 eV, it delocalizes almost completely into weakly bound conduction-band states. This can be expressed as auto-ionization of the compact state:

$$O^{2-} + H^0 \rightarrow OH^- + e_C \tag{17}$$

(and likewise for muonium). Paradoxically, despite its high affinity for protons, i.e. H^+, oxygen has no discernible affinity for *neutral* atomic hydrogen, i.e. H^0, either in ionic or covalent

[14] A good many more muonium hyperfine constants are known (a compendium is given in appendix D) but cannot be plotted in the absence of consistent band-offset information, electron affinity being in principle indeterminate for individual materials: see, for example, Fall *et al* (2002).

oxides. In silica, for example, atomic hydrogen and muonium are chemically stable—the muonium precession signal lives for several microseconds at all temperatures up to at least 1000 K[Note 15]. Atomic muonium and hydrogen are likewise stable (at least on the microsecond timescale, i.e. until they find defects or impurities) in MgO—and indeed in H_2O, ice or water—so that ionicity or polarity is not an issue here. Equation (17) requires the hydrogen or muonium electron to be picked off by the cation, not the anion. For oxides of the heavier transition metals this might not release the electron for conduction, since the charge could be localized by a change of oxidation state:

$$M^{(n+1)+}O^{2-} + H^0 \rightarrow M^{n+} + OH^-. \tag{18}$$

On the other hand, if the electron transfers to empty antibonding states, e.g. the vacant 4s orbitals of Zn^{2+} or 5s orbitals of Cd^{2+}, it is essentially lost to the conduction-band.

The molecular orbital diagrams of figures 23(a) and (b) may give some insight into this behaviour. They imply approach of interstitial H or Mu to the anion and a consequent interaction with the host bonding orbitals (since states at the valence-band maxima have predominantly anion character). This leaves the antibonding (cation) orbitals almost undisturbed and ready to receive an electron into CB states as the level ordering inverts. The opposite assumption of approach to the cation can be considered as favouring formation of the Mu^- or hydride ion, the charge transfer leaving a dangling bond on the anion or a hole in the VB. This latter reaction is hole ionization of the neutral centre and corresponds to the the deep (0/−) acceptor function:

$$H^0 \rightarrow H^- + h_V. \tag{19}$$

Van de Walle and Neugebauer (2003) invoke a charge neutrality argument between cation and anion dangling bonds in support of the commonality of the H(+/−) pinning level, linking it to the origin of the electrochemical scale. It puts their pinning level at −4.5 eV below the vacuum level, somewhat lower than the value of -3 ± 0.4 eV estimated by Kiliç and Zunger (2002) or the muonium thresholds of figures 20–22. The offset is small, but has the curious result that it lies deep in the gap for CdS and CdTe and intersects the valence-band in Ge. In these three instances, therefore, its predictions at first sight appear contrary to the muonium data. Does the muonium pinning level lie higher than that for protium? The possibility of a significant isotope effect is considered and dismissed in the following section. Figure 24 suggests there may in fact be no discrepancy if the criterion for spectroscopic observation of neutral muonium states relates to the position of the (0/+) donor level, relative to the condution-band edge, rather than of the (+/−) pinning level.

10.1. Coexistence of shallow and deep states in CdTe: shallow donor but deep acceptor?

It is significant that CdTe is so close to the threshold, since this is so far the only material in which an apparent *coexistence* of shallow-donor and trapped-atom states has been reported (Gil *et al* 2001). The behaviour is entirely consistent with negative-*U* behaviour, if the (0/+) donor level becomes band resonant but the (−/0) acceptor level remains deep in the gap: this possibility is sketched in figure 24. The rather subtle implication is that, even if a material supports a shallow-donor muonium or hydrogen state, hydrogen will not cause conductivity under equilibrium conditions as long as the (+/−) pinning level remains deep in the gap. GaN may well be another case in point: the calculations of Neugebauer and Van de Walle (1995, 2003) find strongly negative-*U* behaviour for hydrogen in this material, with the (0/+) donor transition lying at or above the conduction-band edge.

[15] ISIS test data (unpublished); see also Dawson *et al* (1997).

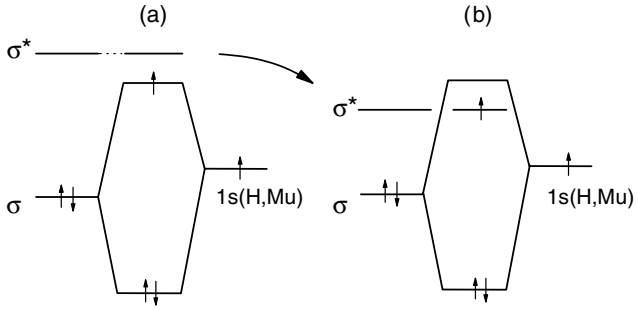

Figure 23. Molecular orbital diagrams for hydrogen or muonium adjacent to the host anion, expressing the transition from (a) deep to (b) shallow donor. In (a) the SOMO is compact; in (b) it has CB character.

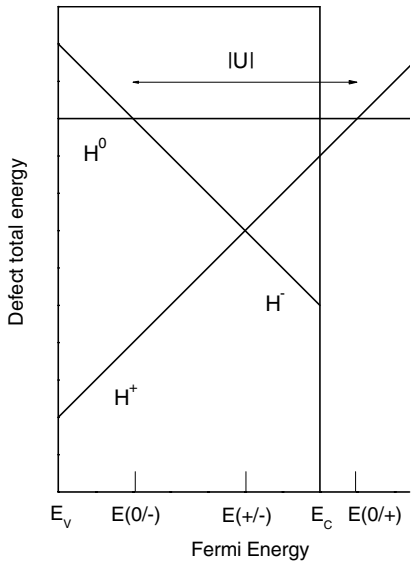

Figure 24. Illustration of the situation in which the (0/+) transition is band resonant, so that the donor becomes shallow, but the (−/0) acceptor level remains deep in the energy gap (for comparison with figure 3). This appears to be the case for CdTe and may well be more widespread, e.g. in GaN (compare figure 1 of Van de Walle and Neugebauer (2003)).

11. Isotope effects: the extrapolation from muonium to hydrogen

Any isotope effect in the (+/−) pinning level would, according to the ionic model of section 9, result from zero-point energy effects in Madelung and polarization energies. Both quantities vary non-linearly with proton or muon position so that different averages are taken by the different spread of their wavefunctions. This harmonic correction is offset by an anharmonic correction, bond-lengths to the anion being typically several per cent longer, and the bond correspondingly weaker, for the muon than for the proton. According to the charge-neutrality argument of Van de Walle and Neugebauer (2003), the pinning level is essentially independent of hydrogen bond-strengths, either to anion or cation, in which case any isotope effect would be negligible.

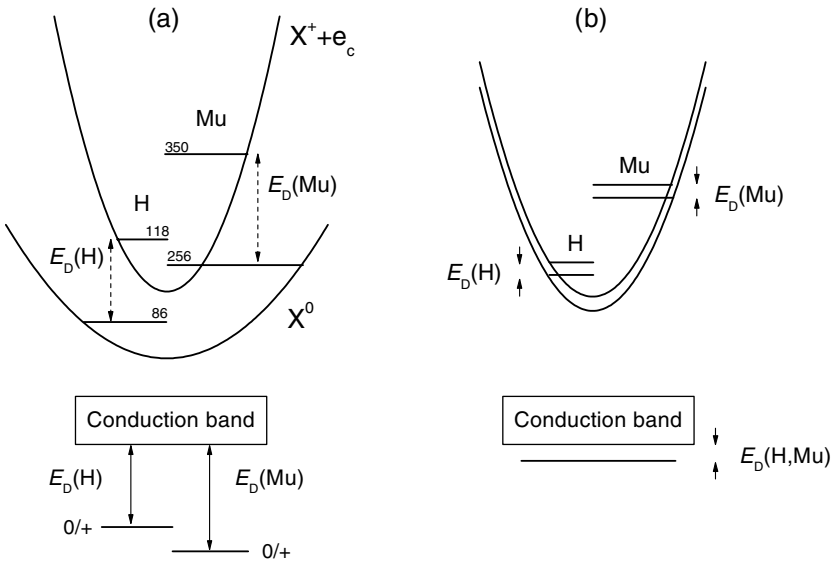

Figure 25. Zero-point energies for muonium and protium, in potential wells drawn for both the neutral centres and positive ions. (The numerical values are in millielectronvolts and are estimated for H_{BC} and Mu_{BC} in Si: see appendix A.) The potential energy for the positive ion includes that of an electron in condution-band, so that the energy differences correspond to ionization energies or donor depths. The isotope shift is significant for deep donors (a) but not for shallow donors (b).

For (0/+) donor levels, the situation is clearer. Isotope effects can be estimated simply and accurately both for deep and shallow donors. Figure 25 shows zero-point vibrational energies for generic muonium and hydrogen centres in their neutral and positive charge states. Harmonic potentials are assumed so that, in each well, the muonium level lies higher than the protium level by a factor $\sqrt{m_{\rm H}/m_{\rm Mu}} = \sqrt{m_{\rm p}/m_\mu} \sim 3$. The zero-point energy for muonium is therefore on a par with the first vibrationally excited level for protium. It can be several tenths of an electronvolt, i.e. comparable with the ionization energy of a deep donor but greater than that of a shallow donor. Of interest is the isotope effect on the ionization energy itself, and therefore on the depth $E_{\rm D}$ of the donor level below the conduction-band minimum. In a harmonic approximation, this is

$$E_{\rm D}({\rm Mu}) - E_{\rm D}({\rm H}) = \frac{\hbar}{2}\frac{(\sqrt{k_+} - \sqrt{k_0})}{\sqrt{m_{\rm p}}}\left(\sqrt{\frac{m_{\rm p}}{m_\mu}} - 1\right) \approx \hbar(\sqrt{k_+} - \sqrt{k_0})/\sqrt{m_{\rm p}} \tag{20}$$

where k_0 and k_+ are the force constants describing the curvature of the two potential surfaces[16]. For the deep donors, the force constant is expected to be smaller for the neutral state than for the positive ion, the antibonding electron softening the potential. An estimate for the deep-donor states of muonium and protium in silicon is compared with experimental data in appendix A. For shallow donors, local screening in the neutral state is minimal, the force constants will be very similar (figure 25(b)) and the expected correction negligible.

[16] Notice that the harmonic frequency is notional: H^0_{BC} in Si would ionize before reaching its first vibrational excited state (Hourahine 2000), for example, as would the shallow donors.

12. Extrapolation to the oxides

Figure 26(a) shows the predictions of density-functional calculations for a subset of oxides due to Kiliç and Zunger (2002). The pinning level of −3.5 eV corresponds closely to the threshold electron affinity of 3.5 eV which accounts for the muonium data of figure 20, although it lies higher than the 4.5 eV calculated by Van de Walle and Neugebauer (2003). Extrapolation to a wider selection of oxides is possible by drawing these candidate thresholds on the band-offset diagrams due to Schmickler and Schultze (1986) and Memming (1983), as in figure 26(b). The correspondence between the vacuum scale and the electrochemical scale is given by these authors; here it is interesting that the proximity of the hydrogen pinning level to the origin of this scale is, according to Van de Walle and Neugebauer (2003), a key to its essentially universal value (an argument which contrasts with but may complement the above ionic model). The implication is that hydrogen could be a cause of electronic conductivity in a surprisingly large number of oxides. Kiliç and Zunger (2002) add Ag_2O, Cr_2O_3, Fe_2O_3, Sb_2O_3, $FeTiO_3$ and $PbTiO_3$ to the list obtained from figure 26. Their inclusion of PbO_2, RuO_2 and IrO_2 is puzzling, since Schmickler and Schultze give only valence-band data for these oxides, noting that they exhibit metallic conductivity. The existence of a bandgap in these materials in fact depends on structure (as discussed for RuO_2 by Hugosson *et al* (2002)); for PbO_2 in the rutile form the bandgap of 4.45 eV given by Madelung (1996) puts its conduction-band edge well above the pinning level. Also puzzling is their inclusion of HgO, for which the cited muonium data indicate a deep rather than a shallow donor (see section 6, above), similar to that of bond-centred hydrogen and muonium in silicon[17]. Since the deep donors are ionized or dissociated at room temperature, however, the distinction is a fine one for most purposes.

Oxides mentioned by Kiliç and Zunger (2002) as being immune to hydrogen doping include HfO_2—an assertion that is challenged by Shluger *et al* (2003) and also by Peacock and Roberston (2003), who find H^+ to be more stable than H^0 in this material, as in ZrO_2. Here it must be noted that Kiliç and Zunger are concerned with the H(+/−) pinning level, whereas Peacock and Robertson calculate the H^0 SOMO energy (approximating to the (0/+) level—see figure 2), so there may be no contradiction, but this point remains to be clarified. Hafnia and zirconia are both amongst the candidates for high-permittivity gate dielectrics (see, for example, Ragnarsson *et al* 2001), so the proximity of their electron affinities to the threshold is crucial here. All these authors agree that sapphire (Al_2O_3) will not exhibit hydrogen-induced conductivity, Peacock and Robertson finding H^- to be the stable charge state in this materal. Here the μSR data are again supportive: a trapped-atom Mu^0 state can be seen whose formation rate, following muon implantation, depends strongly on temperature. An anomalous electric-field dependence of the μSR signals suggests involvement of the Mu^- ion (Brewer *et al* 2000). If confirmed, this is the first example of the deep-acceptor function in oxides. A similar behaviour may be expected in Y_2O_3 (Ragnarsson *et al* 2001).

13. Concluding remarks

The most common neutral defect centres formed by muonium in semiconductors and dielectrics are trapped-atom states, located at the centre of interstitial cages. These are known in over 30 different materials, covalent, polar and ionic, with both tetrahedral and octahedral coordination (see appendix D). The corresponding atomic states of hydrogen have eluded detection in semiconductors (presumably by virtue of their short lifetime against charge-state

[17] Preliminary μSR data indicate a similarly deep-donor, rather than shallow-donor, muonium state in Ag_2O. Magnetic oxides are not suitable for muonium studies, rapid spin relaxation making Mu^0 states difficult to recognize; at the time of writing, μSR studies of the non-magnetic candidates have just begun and will be reported elsewhere.

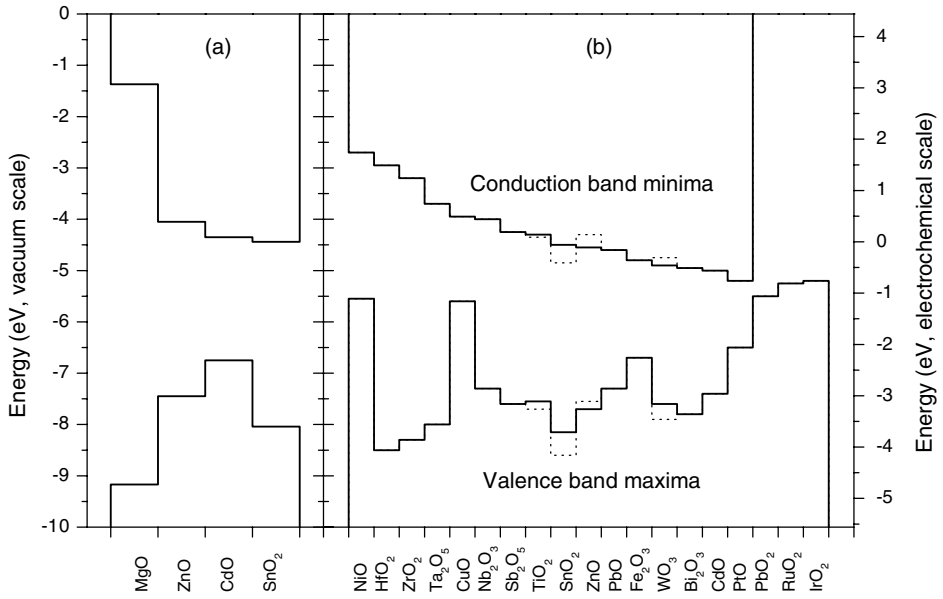

Figure 26. Band-offset diagrams redrawn from (a) Kiliç and Zunger (2002) and (b) Schmickler and Schultze (1986), with dotted additions from Memming (1983). By inference, all materials with CB minima falling below the H(+/−) pinning level are susceptible to hydrogen-induced conductivity.

change or chemical reaction) but are known in a variety of oxides and halides. In terms of a universal model of the electrical activity of hydrogen, it seems likely that these interstitial atoms correspond to the neutral states of deep-level acceptors. Demonstration of the acceptor function, i.e. of the (0/−) charge-state transition, is so far only established with certainty in silicon, where it has been characterized both for muonium and hydrogen (see, for instance, Hitti 1999, Bonde Nielsen *et al* 1999). Similar demonstrations in other materials, together with spectroscopic identification of the resultant negative ion, becomes a high priority for future studies[18].

Contrasting sharply with these compact atomic defects, extended shallow-donor muonium states have now been identified with certainty in five materials—CdS, CdSe, CdTe and ZnO amongst the II–VIs and InN amongst the III–Vs. The analogous shallow-donor state of hydrogen is confirmed by spectroscopic studies in ZnO and suspected from electrical measurements in InN. These are all tetrahedrally coordinated materials, cubic for CdTe (where the shallow donor appears to coexist with a deep trapped-atom state, albeit short lived) and hexagonal for CdS, CdSe, ZnO and InN. It seems likely that other examples will be found amongst materials with high electron affinity, irrespective of structure or bandgap.

In another five of the tetrahedrally coordinated semiconductors (Si, Ge and diamond amongst the group-IV elements, GaAs and GaP amongst the III–V compounds) the trapped-atom (Mu^0_T) state coexists with a deep-donor state located at the bond centre (Mu^0_{BC}). A similar

[18] Likely μSR techniques will be radio-frequency (RF) resonance as a means of analysing for the final muonium state, here the Mu^- ion (see e.g. Hitti (1999); for the pulsed-source variant see Cottrell *et al* (2003)), level crossing resonance when the cation nuclei have quadrupolar nuclei (e.g. Chow *et al* 1995) or RF double resonance when the cation nuclei are only dipolar. When nuclear magnetism is weak, manipulation of the final muonium state with applied electric fields may still be effective (e.g. Brewer *et al* 2000, Eshchenko *et al* 2003b).

deep donor is found in HgO, so far without site determination. A general description of the electrical activity of hydrogen allows for both acceptor and donor levels, as represented in figures 2, 3 or 24. It is tempting to suppose that it is the deep-donor states which evolve and switch to the shallow donors. While shallow-donor states may occupy bond-centre sites, however, this cannot be a requirement—the effective-mass model is not site dependent. Sites antibonding to the anion become favourable in more ionic materials and the necessary stabilization of the bond-centre site by directional bonding (i.e. residual sp^3 character) is absent in octahedrally coordinated materials. Correlation of muonium hyperfine constants with electron affinity (figure 22) suggests instead that it is the cage-centred trapped-atom states which dilate progressively as the conduction-band edge is lowered and undergo a critical transition to the extended state when they become auto-ionizing. The negative acceptor (Mu^-, H^-) is expected to be antibonding to the cation, the ionized or dissociated donor (Mu^+, H^+) antibonding to the anion. This is an easy shift of site across the interstitial cages or channels (figure 4), the neutral intermediary presumably being located more symmetrically.

Whereas ultimately it is the effects on bulk electrical properties that count, the underpinning information on microscopic electronic structure is more widely available for muonium than it is for monatomic hydrogen centres. This is thanks to the sensitivity of μSR spectroscopy to all three charge states, its microsecond timescale and, quite simply, the fact that muons can be implanted into virtually any material. The muonium data on II–VI and III–V semiconductors lend considerable support to the concept of a universal threshold, whereby shallow-donor behaviour is expected in materials with electron affinities in excess of about 3.5 eV. This should not necessarily be identified with the hydrogen pinning level of current theoretical models, however. The criterion for μSR observation of shallow muonium states is band resonance of the (+/0) donor level rather than the (+/−) pinning level. This is true also for ENDOR observations of hydrogen shallow donors. If these spectroscopic observations could be extended to pick up the (0/−) acceptor levels, they would offer a more complete description, giving the elusive Hubbard or Anderson U as well as the pinning level.

The predictions of pinning-level models that hydrogen may be a cause of doping in a wide range of other materials, notably oxides, remains to be assessed. While beneficial to some applications, hydrogen-induced conductivity would be of considerable concern in oxides or nitrides which are under development as thin-film insulators, gate dielectrics or even just compliant layers. Even prime candidates such as hafnia and zirconia are uncomfortably close to the postulated thresholds. Certainly in the initial screening for shallow-donor states in a large number of non-magnetic materials, as a method that does not rely on favourable hydrogen solubility, muonium spectroscopy may be expected to play an important role.

Acknowledgments

The shallow-donor muonium studies have all been the result of collaborative work. I wish to record my indebtedness to all coauthors of the original publications, especially to E A Davis, joint holder of the EPSRC research grant R25361 which funded much of this work, to J M Gil and H Alberto and their students, and to R L Lichti. They have all helped substantially with this review. I thank J J Davies, N Gidopoulos, J S Lord, J Harding, A H Harker, E Roduner and A Weidinger for innumerable discussions and C G Van de Walle, J Robertson, A L Shluger and J Gavartin for early preprints of their (now published) work as well as for valuable suggestions. The constant encouragement of N Ayres de Campos and A M Stoneham has been much appreciated.

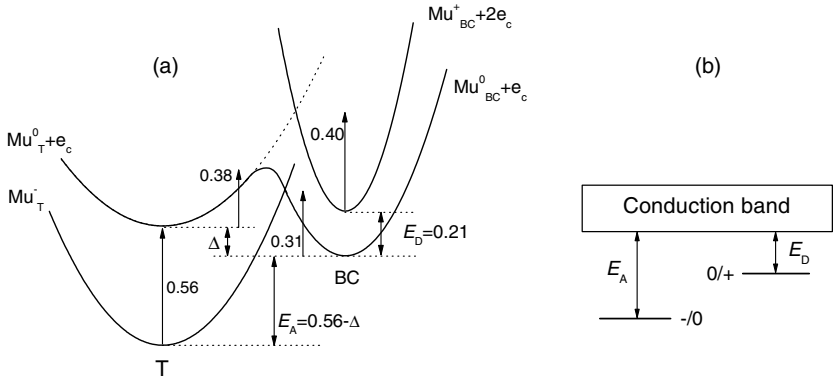

Figure A.1. Configuration diagram for the three charge states of muonium and hydrogen in silicon showing (a) the experimental energy parameters for muonium (in eV, from Hitti 1999) and (b) the correspondence with the deep donor and acceptor levels (from Cox 1995).

Appendix A. Deep-level amphoteric muonium and hydrogen

A.1. Tribute to T L Estle

Figure A.1 illustrates the various sites and charge states for muonium in silicon. Just four states are sufficient to account for the great majority of spectroscopic and dynamical μSR data, namely Mu^{+}_{BC}, Mu^{0}_{BC}, Mu^{0}_{T} and Mu^{-}_{T}, linked by a relatively simple set of potential energy wells and barriers. A huge volume of spectroscopic and dynamical muonium data, accumulated by the μSR community over three decades, is condensed here into a deceptively simple interplay of the three charge states amongst two crystallographic sites in diamond-type or zincblende structures. This interpretation stands especially as a tribute to the persistence and vision of the late T L Estle, to whose memory this review of the new discoveries is dedicated.

Discovery of the metastability of the neutral centre dates from the earliest μSR studies (Brewer *et al* 1973) although it was some time before the importance of the BC site was appreciated and confirmed (Cox and Symons 1986, Kiefl *et al* 1987, 1988). The concept of bistability of the charged centres, as well as their more pronounced site preferences, came largely from *ab initio* calculations for hydrogen: see, for example, Van de Walle *et al* (1988). Without the hyperfine signatures of the paramagnetic neutral states, site determination for Mu^{+} and Mu^{-} is virtually impossible in Si, but becomes possible in GaAs thanks to the abundant quadrupolar nuclei in III–V compounds (Chow *et al* 1995, Chow 2003). Schemes essentially similar to figure A.1 are believed to apply to muonium in the other group-IV semiconductors Ge and diamond and also to the III–V compounds GaAs and GaP. The relative stability of Mu^{0}_{BC} and Mu^{0}_{T} is less pronounced in Ge (Lichti *et al* 1999) and possibly inverted in the other materials (Estreicher 1994). In the more ionic III–Vs, the positive ion may also be metastable, with the anion cage centre of the T_V site competing with the bond centre (Lichti *et al* 1997, Lichti 2003a).

Whilst the hydrogen analogue of Mu^{0}_{BC} in silicon (i.e. H^{0}_{BC}, originally dubbed the AA9 centre) is detectable by ESR, albeit after illumination of the sample (Gorelkinskii and Neninnyi 1987, 1991, Bech Nielsen *et al* 1994), that of Mu^{0}_{T} is not. Herring *et al* (2001) question the importance of H^{0}_{T}, arguing that the adiabatic potential energy surface for H^{0} is downhill all the way from T to BC, with no barrier. This amounts to treating the proton classically, or as infinitely heavy. The T to BC barrier could in part be kinetic, the heavy Si atoms being

unable to respond adiabatically to the instantaneous position of either H or Mu. Porter *et al* (1999) tackle this problem with a double adiabatic approximation, in which the electron, muon or proton and lattice degrees of freedom are separately decoupled (curiously, however, these authors find Mu_T^0 to be more stable than Mu_{BC}^0). Other theoretical studies support the concept of a T to BC barrier, and consequent H^0 metastability: bibliographies are given, for instance, by Seager *et al* (1995) and Jones *et al* (2000). Since Mu_T^0 diffuses rapidly at all temperatures, it seems likely that H_T^0 does likewise, reacting with other defects and impurities before it can be detected in ESR studies. This is the view of Bonde Nielsen *et al* (1999), who identify H_T^0 as the elusive transport state of interstitial hydrogen in silicon.

Figure A.1(b) shows how the correspondence is made with the donor and acceptor levels of figures 2 and 3. It is easy to see how $U = E_A - E_D$ can be either positive or negative, in view of the site-change involved, although deductions from experimental data have been surprisingly contentious (see, for instance, Seager *et al* 1995). A common error is to identify the acceptor depth E_A with the energy difference between H_T^- and H_T^0 rather than between H_T^- and H_{BC}^0. The former is spectroscopically the more accessible transition (the single-site ionization energy of the hydride ion) and is undoubtedly involved in second-electron capture and release but the latter (between the negative ion and the stable neutral) is the thermodynamic definition (Cox 1995). Tantalizingly, although the majority of the energy differences and barriers of figure A.1 have been determined for muonium (Kreitzman *et al* 1995, Lichti 1995, Hitti 1999), the notable exception is the energy difference Δ between Mu_T^0 and Mu_{BC}^0, which is crucial to determining the sign of U. Hitti (1999) note that U is negative only if Δ is less than 0.35 eV.

A.2. Isotope effects

The close correspondence of all the parameters of figure A.1(a) for muonium with those for hydrogen is documented by Kreitzman *et al* (1995) and updated by Bonde Nielsen *et al* (1999); it entirely validates the use of the muonium parameters as a guide to the electrical activity of hydrogen. The point is that the donor and acceptor depths of figures 2 and 3 correspond to energy differences in figure A.1 so that, even though the zero-point energies can be a substantial fraction of an electronvolt, isotope effects in the electrically active levels are not large. This is detailed in figure 25 for donor states. Identical potential-energy surfaces are assumed for the two isotopes. The zero-point energy for muonium is greater than that for protium in the same well by a factor of $\sqrt{m_H/m_{Mu}} \sim 3$. The difference between the muonium and hydrogen donor depths—210 meV for $Mu_{BC}^{0/+}$ as against 175 meV for $H_{BC}^{0/+}$—is closely accounted for in terms of the different zero-point energies for muon and proton and the rather softer harmonic potential for the neutral centre, as opposed to the positive ion. Here the zero-point energies, entered as small numerals in figure 25, have been estimated using force constants (vibrational frequencies) due to Hourahine (2000).

A.3. Which state evolves to the shallow donor?

Figure A.1 serves also to illustrate the fact that the neutral donor in silicon is *not* the trapped-atom state H_T^0; instead, it is the bond-centred species H_{BC}^0 whose singly occupied orbital borrows anti-bonding character from the host lattice (Cox and Symons 1986, see also figure 9(c)), guaranteeing its donor qualities. It might reasonably be expected that this is this state which dilates further and evolves into the shallow-donor state in other materials. The trapped-atom state Mu_T^0 seems more readily associated with the acceptor function, converting by second electron capture to Mu_T^-. Yet according to figure 22, dilation of Mu_T^0 correlates with the host electron affinity and seems to undergo the critical transition to the extended state.

Appendix B. Hyperfine spectroscopy

A more formal account of muonium spectroscopy is given in this appendix, collecting together expressions commonly used in the data analysis. For the transverse-field spectroscopy, formulae for the precession frequencies carry over with some judicious selection from those used in electron–nuclear double resonance (ENDOR): see, for example, Abragam and Bleaney (1970), Pake and Estle (1973), Slichter (1978) or Atherton (1993). Compendia specific to muonium spectroscopy such as those given by Patterson (1988) and Senba (2001) are useful, however, and many of the zero-field and longitudinal-field considerations are unique to μSR.

B.1. Isotropic muonium: transverse-field spectroscopy

As a starting point, consider atomic muonium in a static magnetic field: the spin Hamiltonian takes the form—in units of frequency—

$$\mathcal{H}/h = \nu_e S_z - \nu_\mu I_z + A\mathbf{I}\cdot\mathbf{S}. \tag{21}$$

The free-atom hyperfine constant is $A = 4463$ GHz in its 1s ground state but, with other values of A (normally lower, see figures 22 and D.1), equation (21) applies equally to all isotropic muonium defect centres, wherever couplings to nuclear spins of the host material can be neglected. Here $\mathbf{I}$ and $\mathbf{S}$ are the muon and electron spins (both one-half); $\nu_\mu = (1/2\pi)\gamma_\mu B$ and $\nu_e = (1/2\pi)(g/2)\gamma_e B$ are their respective Larmor frequencies in the field B, whose direction defines the z-axis. For the trapped-atom muonium states, electron spin–orbit coupling is negligible so that $g = 2$, as for the free atom. The gyromagnetic ratios are $\gamma_\mu = 2\pi \times 136$ kHz mT^{-1} and $\gamma_\mu = 2\pi \times 28$ MHz mT^{-1}. (Beware the factors of 2π! Note also that $\gamma_\mu/\gamma_e = m_\mu/m_e \sim 1/200$, since muons and electrons both belong to the lepton class of particles: see table 1.) The eigenvalues, originally given for atomic hydrogen by Breit (1929), define a generic energy-level diagram as sketched in figure B.1(a). In transverse field, i.e. with the initial muon spin polarization along x or y, the muon spin rotation signal contains the four precession frequencies ν_{12}, ν_{23}, ν_{14} and ν_{34}. These have equal intensity in the limit of low field, and the so-called triplet precession signals at ν_{12} and ν_{23} are commonly used to identify atomic muonium. A more precise measurement of hyperfine constant is obtained in the high-field or Paschen–Back regime where only ν_{12} and ν_{34} are visible, given by

$$\nu_{12,34} \approx |\nu_\mu + A^2/4\nu_e \mp A/2|. \tag{22}$$

An exact and direct measurement of the hyperfine constant A is then given by the sum or difference of the two frequencies, according to whether $A/2$ is greater or less than $\nu_\mu + A^2/4\nu_e$ (the quadratic term is usually small, and the appropriate condition is easily recognizable by taking spectra at several fields).

B.2. Anisotropic muonium: transverse-field spectroscopy

For the donor states of muonium, deep and shallow, equation (21) must be generalized to allow for hyperfine anisotropy and for electronic g-values other than two. Axial symmetry for the hyperfine tensor and and isotropic g-values are usually sufficient, giving additional splittings or avoided crossings in the energy-level diagram, as in figure B.1(b). The Hamiltonian is then cumbersome for arbitrary orientation of the tensor but, again in a high-field regime where the electron Zeeman energy is the dominant term of the spin Hamiltonian, can be written (Slichter 1978)

$$\mathcal{H}/h = \nu_e S_z - \nu_\mu I_z + (A_\perp \sin^2\theta + A_\parallel \cos^2\theta) I_z S_z + (A_\perp - A_\parallel)\sin\theta\cos\theta I_x S_z. \tag{23}$$

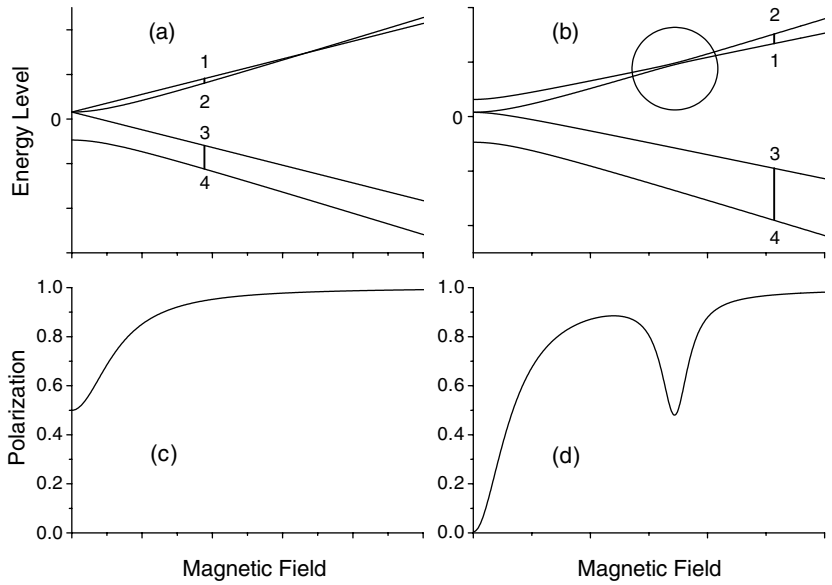

Figure B.1. Breit–Rabi diagram for atomic muonium or for isotropic muonium defect centres (a) and its modification for small axial anisotropy (b), here drawn for the symmetry axis nearly perpendicular to the magnetic field. The corresponding repolarization or decoupling curves are shown in (c) and (d), respectively, with the resonance in (d) corresponding to the anticrossing circled in (b). (Simulations due to Lord. A fictitious electron g-value of 0.025—or, equivalently, a fictious ratio $\gamma_e/\gamma_\mu = 5$—has been used to generate curves of reasonable proportions.)

Here θ is the angle between the magnetic field (the z-axis) and the axis of symmetry of the defect centre, i.e. the principal axis of the hyperfine tensor. The two hyperfine terms in equation (23) are commonly interpreted as generating an effective field, as experienced by the muon spin, with components both parallel and perpendicular to the externally applied field. The relevant frequencies become (see, for example, Hintermann *et al* 1980)

$$\nu_{12,34} = \{[\nu_\mu \mp \tfrac{1}{2}(A_\perp \sin^2\theta + A_\parallel \cos^2\theta)]^2 + \tfrac{1}{4}(A_\perp - A_\parallel)^2 \sin^2\theta \cos^2\theta\}^{1/2} \tag{24}$$

and are invariably observed as satellite lines positioned almost symmetrically about ν_μ, the muon Larmor frequency. No transition at ν_μ figures in the Mu^0 Breit–Rabi diagram, of course, although in experimental spectra a line at this frequency is usually observed from Mu^+ states, either in the sample or the sample mount. Detection of Mu^0 spectra as two symmetric satellites to the Mu^+ line, as in figure 6, corresponds to a second level of high-field approximation where the muon Zeeman energy now dominates the hyperfine terms in the Hamiltonian. An approximate expression for the satellite separation,

$$\nu_{34} - \nu_{12} \approx A_\perp \sin^2\theta + A_\parallel \cos^2\theta, \tag{25}$$

given by the parallel hyperfine field, works well for the range of parameters encountered in the shallow-donor states. In this double high-field limit, the quadrature term makes a quite negligible correction. Equation (25) is certainly a better approximation than the quantity $\{A_\perp^2 \sin^2\theta + A_\parallel^2 \cos^2\theta\}^{1/2}$, which also appears in the literature as a measure of hyperfine field. This latter expression—as used by Shimomura *et al* (2002) to interpret shallow-donor muonium data from ZnO single crystals—is more appropriate to the intermediate-field regime where muon Zeeman energy is small compared with the hyperfine parameters, although it is doubtful if the two forms could be distinguished in experimental data.

Table B.1. Hyperfine parameters for the deep-donor states. (Those for the shallow donor in ZnO are also shown for comparison. The distance d is the effective electron–muon separation corresponding to the dipolar parameter D, for pointlike localization of one Bohr magneton.)

	References	$A_\parallel$	$A_\perp$	$A_{\rm iso}$ (MHz)	D (MHz)	d (nm)
HgO	Cox *et al* (2001a)	20	18	15	5	0.5
Si	Cox and Symons (1986)	−16	−93	−67	51	0.24
Ge	Cox and Symons (1986)	−17	−131	−96	69	0.25
GaAs	Cox and Symons (1986)	218	89	126	94	0.25
GaP	Cox and Symons (1986)	219	79	132	86	0.24
(ZnO)	Alberto *et al* (2001)	0.80	0.33	0.49	0.31	1.3

Table B.2. Hyperfine parameters for shallow-donor muonium in zinc oxide, comparing the results of different experimental methods, transverse field (TF) or longitudinal field (LF) on single-crystal and powder samples. (The bottom line refers to a different centre identified by Shimomura *et al*: see section 4.)

Technique	$A_{\rm iso}$ (kHz)	D (kHz)	References
Powder TF	500 ± 20	260 ± 20	Cox *et al* (2001a)
Single-crystal TF	490 ± 10	310 ± 10	Alberto *et al* (2001)
Powder TF	506 ± 3	270 ± 3	Alberto *et al* (2001)
Powder LF	457 ± 4	267 ± 6	Alberto *et al* (2001)
Single-crystal TF	491 ± 14	265 ± 14	Shimomura *et al* (2002)
Single-crystal TF	293 ± 20	286 ± 19	Shimomura *et al* (2002)

Following Cox and Symons (1986), it is useful to decompose the principal components of the hyperfine tensor into an isotropic or contact term which is a measure of (unpaired electron) spin density at the muon site, together with a traceless dipolar term which depends on how the overall spin density is distributed in the muon's vicinity:

$$A_{\rm iso} = \tfrac{1}{3}\{A_\parallel + 2A_\perp\}, \tag{26}$$

$$D = \tfrac{2}{3}\{A_\parallel - A_\perp\}. \tag{27}$$

In terms of these parameters, the high-field satellite splitting of equation (25) becomes

$$\nu_{34} - \nu_{12} \approx A_{\rm iso} + \frac{D}{2}(3\cos^2\theta - 1), \tag{28}$$

which is the form used by Gil *et al* (2001). Values of all the various parameters, as determined from single-crystal studies, are given in table 4. We return to the interpretation of the contact interaction $A_{\rm iso}$ at the end of this appendix. As a guide to the magnitude of dipolar interactions in the shallow-donor states, the value of $D = 91$ kHz for CdS is the same as would be generated by one Bohr magneton localized at distance 2 nm from the muon. This is to be compared with the effective Bohr radius of $a^* = 25\ a_0 = 1.25$ nm in the envelope-function model of figure 9(a). If the extended wavefunction were spherically symmetric about the muon site, D would be vanishingly small; it is the absence of exact symmetry, the small p-wave component of the band or Wannier functions and their cross-terms with 1s(Mu) which give the slight dipolar character. In contrast, for the deep-donor state $\mathrm{Mu}^0_{\rm BC}$ in silicon, the value $D = 51$ MHz corresponds to the unpaired electron effectively localized about 0.2 nm from the muon site: see table B.1.

B.3. Powder-pattern spectra

Superposition of satellite lines with their polycrystalline weightings gives distinctive powder-pattern lineshapes. The basic spectral density function (Alberto *et al* 2001) is

$$\frac{\mathrm{d}P}{\mathrm{d}\nu} = \frac{\mathrm{d}P}{\mathrm{d}\theta}\,\frac{\mathrm{d}\theta}{\mathrm{d}\nu} = (3D/2)^{-1/2}(2|\nu - \nu_0| - A_{\mathrm{iso}}/2 + D/2)^{-1/2}, \tag{29}$$

for $\nu_0 - A_\parallel/2 \leqslant \nu \leqslant \nu_0 - A_\perp/2$ and $\nu_0 + A_\perp/2 \leqslant \nu \leqslant \nu_0 + A_\parallel/2$, where $\nu_0 \approx \nu_\mu$ is the centre of the distribution. This is the form superimposed on the experimental spectra for powder samples in figures 10 and 16 (ZnO and HgO). The inner splitting is seen to correspond to $A_\perp$, with peak intensity corresponding to the most probable value of θ; the outer splitting corresponds to $A_\parallel$. The two principal values of the hyperfine tensor are therefore readily obtained from a single powder-pattern spectrum. Equation (29) can also be convolved with a narrow Gaussian to round off the singularities: Alberto *et al* (2001) use such a form to fit the experimental polycrystalline lineshapes, finding values for A_{iso} and D which are indistinguishable from the results of a more painstaking study of the orientation dependence of single-crystal spectra. Table B.2 compares the results, together with those from zero-field and longitudinal-field spectroscopy, described below.

B.4. Zero-field and longitudinal-field spectroscopy

In zero or longitudinal field, oscillations of the muon polarization are observable at frequencies corresponding to transitions between those energy levels of figures B.1(a) and (b) which are not linear in field. Here longitudinal means in line with the initial muon polarization, i.e. along z. Curvature of the energy levels implies a mixing of the electron and muon spin states and a consequent oscillatory exchange of polarization. Unlike the transverse-field signals, there is no sense of precession to these oscillations if the sample is polycrystalline or if the hyperfine interaction is purely isotropic—they are then invisible in the transverse directions, x or y. In the isotropic case, the zero-field frequency is a direct measure of hyperfine constant. For anisotropic centres the additional frequencies may also be visible in favourable cases, but are often damped or obscured by nuclear couplings. In the region of the avoided level crossing another relatively strong oscillation appears, visible in longitudinal field. The nuclei are then largely decoupled and give less interference. There is also a precise value of the field, roughly central to the avoided level crossing, for which the oscillation frequency coincides for all possible orientations: it is then equal to $3/8D$ and so gives a direct measure of the dipolar parameter. This condition allows detection in polycrystalline samples and was dubbed by Patterson (1988) the 'magic field'. The signal for ZnO is shown in figure B.2. A fitting of these zero- and longitudinal-field μSR spectra for ZnO due to Alberto *et al* (2001) yields hyperfine parameters that are essentially identical with those from the transverse-field precession signals: see table B.2.

B.5. Repolarization

Underlying the oscillatory signals seen in zero and longitudinal fields is a time-average muon polarization $2\langle I_z\rangle$ which is readily determined from the forward–backward asymmetry in the muon decay. In fact, since its measurement requires no timing resolution in the positron counting, much higher data rates can be achieved than for oscillatory signals, using integral-counting methods when large muon fluxes are available. This is something of a μSR speciality. Referring first to the Hamiltonian of equation (21), as a field is applied and increased in this *longitudinal* direction, the oscillation frequency increases as ν_{24}—the muon–electron flip–flop

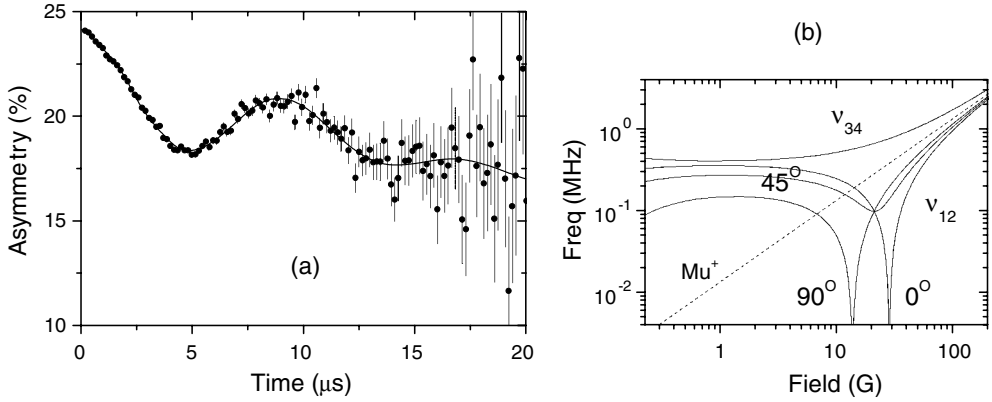

Figure B.2. Longitudinal-field response of the muon spin at the so-called magic field for muonium in a powder sample of ZnO, where the transition frequency ν_{12} coincides for all orientations (Alberto *et al* 2001).

frequency—but its amplitude decreases and the time-average polarization increases according to

$$2\langle I_z\rangle = \frac{1}{2} + \frac{1}{2}\frac{B^2}{B^2 + B_0^2}. \tag{30}$$

This is the form of the curve in figure B.1(c); the characteristic hyperfine field is itself a measure of the hyperfine constant:

$$B_0 = 2\pi A/(\gamma_e - \gamma_\mu) \approx 2\pi A/\gamma_e. \tag{31}$$

Such field-dependent recoveries of polarization are variously known as decoupling, repolarization or quenching curves. For free or vacuum state muonium, $B_0 = 160$ mT; for the shallow-donor states it is only 10–20 μT although the repolarization is then stretched towards 1 mT, dominated by decoupling of the distributed nuclear couplings. This effect is more clearly seen in the time domain as an effective cross-relaxation rate: see figure B.4(b).

Neutral muonium states may often be revealed and characterized by such longitudinal-field repolarization curves, even when transverse-field precession signals or zero-field oscillations are invisible, e.g. when the muonium is formed slowly or when, in contrast, it is a short-lived precursor of some other final state. Coexistence of the deep trapped-atom state with the shallow donor in CdTe was demonstrated in this way (Gil *et al* 2001), recovering the missing polarization of figure 14(b). Fitting equation (30) to the repolarization curves then gives an estimate of the hyperfine constant, although this is at best imprecise and at worst falsified by nuclear dipole couplings or dynamical effects.

B.6. Level crossing resonance

Anisotropy of the coupling results in additional loss of polarization at low field, as shown in figure B.1(d). (Approximate formulae are given by Pratt (1997).) More dramatically, it also results in a resonant loss of polarization at the avoided level crossing. One can say that the applied field tunes out the longitudinal component of hyperfine field at this resonance (close to $B = \pi A_\perp/\gamma_\mu$ for polycrystalline samples), leaving the muon polarization free to precess about the transverse components (equation (23)). Data for the anisotropic muonium centre in HgO provide a particularly striking example, shown in figure B.3.

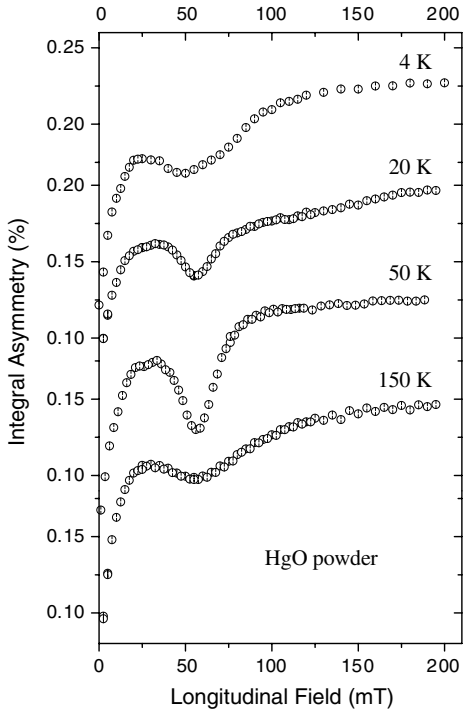

Figure B.3. Level crossing resonance in HgO. Resonant loss of muon polarization occurs in a longitudinal magnetic field corresponding to the anticrossing circled in figure B.1(b): compare figure B.1(d). The time-average polarization is shown as a function of longitudinal magnetic field for various temperatures, the changes in line-shape indicating the onset of diffusional motion. The resonance persists to high temperatures, however, indicating that the hyperfine anisotropy is not fully averaged by the motion.

It is quite remarkable that an anisotropic centre may be observed and its hyperfine parameters determined in a polycrystalline sample, using a longitudinal magnetic field and minimal timing resolution. This is surely unique to μSR spectroscopy. The equivalent resonances are well known for Mu^0_{BC} in silicon, both crystalline and polycrystalline (Patterson 1988, Cooke 1994); they even provide a glimpse of the bond-centred state in amorphous silicon (Davis *et al* 1995, Davis 1996). The resonance relies on avoidance of the level crossing and so is absent for isotropic muonium states, or for defect centres where anisotropy is rapidly and spherically averaged by diffusional motion.

B.7. Nuclear or superhyperfine interactions

Resonant cross relaxation from the muon to surrounding nuclear spins is in principle possible, mediated by their separate hyperfine couplings to the electron, whether these couplings are isotropic or not. Sharp resonances of this sort are well known for muoniated organic radicals, where they permit a mapping of the spin density distribution throughout the molecule. A particularly striking example is the determination of ^{13}C couplings in the muonium adduct of C_{60} (Percival *et al* 1995). The equivalent mapping for Mu^0_{BC} centres in GaAs and GaP, as well as in Si using the sparse ^{29}Si couplings, confirmed their BC location and electronic structure

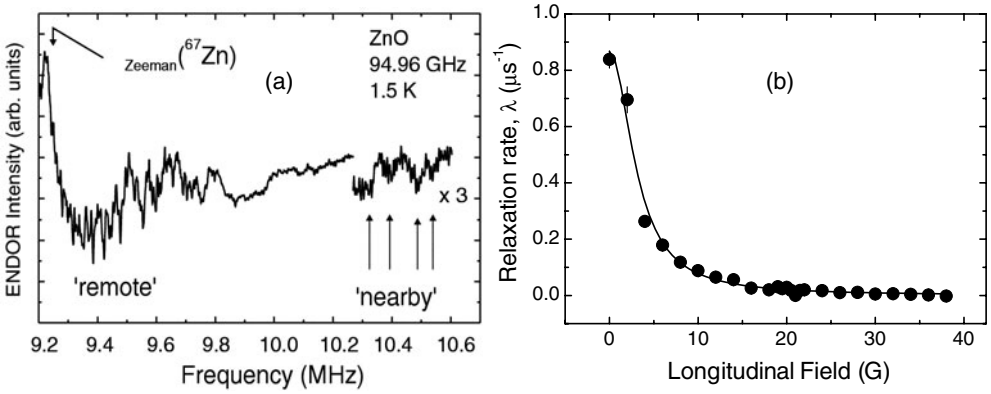

Figure B.4. ^{67}Zn superhyperfine interactions in ZnO. In Zn–ENDOR spectra for the hydrogen centres (a), individual couplings are just resolved. (Reproduced with permission from Hofmann *et al* (2002), copyright (2002) American Physical Society.) They are not resolved in μSR spectra but are the likely cause of the relaxation of the muon spin visible in low longitudinal field, easily decoupled (b). Here λ is the exponential cross-relaxation rate or inverse relaxation time (Alberto *et al* 2001).

(Kiefl *et al* 1987, 1988). The topic has been reviewed by Kiefl and Kreitzman (1992), and the theory of the resonance lineshapes developed by Kreitzman and Roduner (1995).

The Zn-ENDOR spectrum of figure B.4(a) reveals ^{67}Zn superhyperfine couplings up to a detected maximum of about 3 MHz. These are believed to correspond to Zn nuclei immediately adjacent to shallow-donor hydrogen centres in ZnO (Hofmann *et al* 2002). For their muonium counterparts, it has not yet proved possible to resolve individual ^{67}Zn couplings, whether by muon spin rotation or level crossing resonance. However, in longitudinal magnetic field, an overall muon spin relaxation can be discerned, whose suppression with increasing field must represent decoupling of these superhyperfine interactions, as in figure B.4(b). As a rough estimate, a decoupling field B_d of several millitesla defines an interaction strength of $(1/2\pi)\gamma_\mathrm{e}B_\mathrm{d} \sim 5$ MHz. The field dependence is not easily inverted to give the distribution of couplings, but a comparison with simulations is given by Lord *et al* (2001).

These values are about an order of magnitude greater than is expected from the model of figure 9(a), where a single envelope function describes the unpaired electron density on both the central muon and the surrounding Zn nuclei. Hofmann *et al* (2002) report a similar difficulty in relating the proton and ^{67}Zn nuclear couplings. The free-ion Zn coupling is given by Morton and Preston (1978) as 2 GHz and should in this model be reduced by a factor $(A_\mathrm{iso}/A_0)\exp(-2r/a^*)^2 \geqslant 10^4$, i.e. should not exceed 0.2 MHz. This implies that the unpaired electron density may not, in fact, be maximal at the central muon or proton. The radii a^* of the envelope functions presented in table 4 may be somewhat overestimated in consequence. Nonetheless, the assumption that the contact interaction on the muon the scales as $1/a^{*3}$ remains reasonable; together with equations (4) and (5) this leads to the following relation with binding energy:

$$R^* = \frac{R_y}{\varepsilon}\left(\frac{A_\mathrm{iso}}{A_0}\right)^{1/3}, \tag{32}$$

which may be seen to be approximately borne out by the data of table 5.

Appendix C. Electron capture, ionization and occupation statistics

C.1. Muonium formation

Since muons are implanted as the positive ions, neutral muonium must be formed by electron capture at some stage of their slowing down and eventual thermalization. The process does not rely on any pre-existing conduction electron concentration—the best yield and clearest signals are usually seen in undoped or lightly doped material. In contrast, the hyperfine spectra of the paramagnetic states are washed out by spin exchange or second-electron capture in heavily doped n-type material. Neutral muonium centres are also formed and seen as long-lived states—on the microsecond timescale of μSR spectroscopy at least—even when hydrogen is expected to be a negative-U centre, so that the neutral states should not be thermodynamically stable. Of course, the muonium centres cannot equilibrate amongst themselves. In many experiments there is no more than one muon in the sample at a time. Even using a pulsed muon source, or a continuous source in high-flux, integral counting mode, they are always in extreme dilution and the probability of two Mu^0 centres interacting is quite negligible.

To date, all muonium studies have involved implantation of muons with an incident kinetic energy of 4 MeV or greater. (Only recently have moderation techniques been developed to the point where muonium formation might soon be studied in the absence of accompanying radiolysis: Morenzoni *et al* (2003).) As the initially energetic muons slow down, electrons may be acquired epithermally, stripped from either host or defect atoms. In fact, the muonium electron can be stripped in turn, so that the final stages of thermalization are believed to involve a succession of $Mu^+ \rightleftharpoons Mu^0$ charge-exchange cycles as well as the ejection of electron–hole pairs (Brewer *et al* 1975, Senba 1990). This defines a prompt partition into Mu^0 and Mu^+ fractions, as the energy drops below threshold for the charge-exchange resonance.

Even muons that thermalize as Mu^+, however, have a second chance to acquire electrons from their own radiolytic track. This latter process is more akin to the normal processes of carrier capture by Coulombic defect centres and there is increasing evidence of its importance in controlling the observed neutral fraction for the donor states of muonium, both deep and shallow, if not for the trapped-atom states. This comes from studies involving the manipulation of Mu^0 and Mu^+ yields with externally applied electric fields, e.g. for the bond-centred or deep-donor muonium states in Si, GaAs and GaP (Storchak *et al* 1997, Eshchenko *et al* 2003a) as well as for the shallow-donor state in CdS (Eshchenko *et al* 2003b). The capture range is found to be typically 10^{-6}–10^{-5} cm, taken to be the distance between the thermalized muon and the last ionization or charge-exchange events. Quite modest electric fields of a few kV cm^{-1} are sufficient to suppress the muonium yield completely. For the deep donor in GaAs this is quite insufficient to field-ionize the ground state of Mu^0_{BC} and so is taken as evidence that the capture proceeds through weakly bound excited states. For the shallow donor in CdS, the data are shown in figure C.1; the critical electric field seems to correspond to field-ionization of the first excited ($n = 2$) state of the hydrogenic series, suggesting this to be a bottleneck to subsequent cascade to the ground state. In the following section we consider whether a metastable $n = 2$ state could be the cause of the puzzling hyperfine anomaly in ZnO.

C.2. Metastable excited states?

With a comprehensive ESR and ENDOR study of hydrogen itself now published (Hofmann *et al* 2002), a comparison with the μSR parameters for muonium becomes possible. This is the first such comparison for a shallow-donor state. It is of similar importance, therefore, to the validation of Mu^0_{BC} as a model for the deep-donor H^0_{BC} or AA9 centre in silicon (appendix A). The proton-ENDOR spectrum, reproduced in figure 12, shows a pair of lines centred on

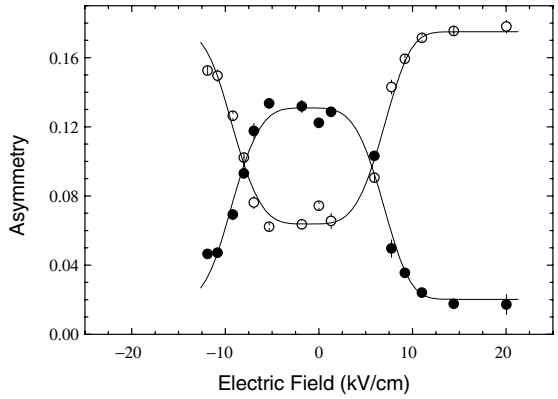

Figure C.1. Electric field dependence of the Mu^+ (open circles) and Mu^0 (filled circles) asymmetries in polycrystalline CdS at 11 K (reproduced with permission from Eshchenko *et al* 2003b: copyright (2003) American Physical Society).

the proton Larmor frequency in the same manner that hyperfine satellites are seen in a μSR spectrum. (In conventional ENDOR the opposite signs of the two satellite lines confirm positive spin density on the proton, as assumed for the muon in figure 9(a).) The proton and muon hyperfine splittings are not in the expected ratio, however. Expressed in units of frequency, they should be related—for identical electron spin densities—by the ratio of the proton and muon magnetic moments (table 1):

$$A_{\rm iso}^{\rm H}/A_{\rm iso}^{\rm Mu} = \mu_{\rm p}/\mu_\mu = \gamma_{\rm p}/\gamma_\mu \approx 0.3. \tag{33}$$

Hofmann *et al* (2002) do not extract the contact interaction or give details of the orientation dependence of the proton splitting but their c-axis ($\theta = 0$) splitting of 1.4 MHz is almost twice the muonium value of $A_\parallel = A_{\rm iso} + D = 0.76$ MHz (Cox *et al* 2001b). Assuming $A_{\rm iso}$ and D to scale in the same manner, this seems to imply a spin density on the proton some six times greater than that on the muon. There is no precedent for an isotope effect of this size in muon and proton hyperfine constants. In view of the close similarity between the other properties of the muonium and protium donors, this anomaly is surprising.

The possibility that the shallow Mu^0 state seen in ZnO does not reach its $n = 1$ ground state within the muon lifetime is discussed by Cox (2003b). For an electron stranded in the $n = 2$ state (the principal quantum number referring not to the compact or free muonium atom but to the effective-mass or shallow hydrogenic series), its density at the nucleus or charge-defect site is reduced by a factor $1/n^3 = 1/8$. This would largely explain the hyperfine anomaly but would introduce instead a discrepancy in binding energy, which varies as $1/n^2$. Despite the uncertainty in the muonium value (19–58 meV in figure 11 and table 5), it is certainly not four times smaller than the hydrogen donor activation energy, given by Hofmann *et al* (2002) as 35 ± 5 meV. Van de Walle (2003) suggests that the discrepancy is a question of different crystallographic sites reached by implanted muons and grown-in hydrogen impurity, i.e. that it has to do with the different manner and timescale of their incorporation. Although the effective mass model in its simplest form is not expected to be site dependent, it could also be that the Wannier functions vary sufficiently rapidly for the muon and proton to take such different zero-point averages of the contact interaction. Could this in fact be the difference between the BC and the AB site—the one being in much closer proximity to a node in the electron wavefunction than the other? It remains to be considered whether metastability within the

$n = 1$ state, due to valley–orbit splittings, or the site dependence of these interactions, could be responsible.

C.3. Muonium ionization: estimation of the donor depth

The ionization regime presents an unusual balance between electron capture and loss, which again might be out of equilibrium. If the electron is acquired epithermally, i.e. irrespective of lattice temperature, it will be lost on thermalization with a temperature-dependent ionization rate, this latter measurable directly as a damping of the μSR signal, i.e. the linewidth of the frequency spectrum. In such a case, μSR sees only the approach to equilibrium. Referring to figures 7 and 19, however, the shallow-donor Mu^0 signals are seen to disappear as much by loss of amplitude as by lifetime broadening. This is rather unusual and is an additional indication of a temperature-dependent formation mechanism. It is actually somewhat reassuring, implying that the measured ionization energies are indeed characteristic of the muonium centres themselves: the satellite lines are not simply washed out by spin exchange when native donors ionize.

The exact complementarity between the Mu^+ and Mu^0 fractions, accounting for the full incoming polarization, even suggests an equilibrium balance between the two charge states. In that case, the standard expression for donor dissociation can be adapted, as in the following treatment due to Davis (2001):

$$\frac{n(N_A + n)}{(N_D - N_A - n)} = \frac{N_c}{2} \exp\left(-\frac{E_D}{kT}\right). \tag{34}$$

Here n is conduction electron density, N_D the total number of shallow donors, N_A the number of shallow acceptors and the equation applies to the case of partial compensation, $N_D/N_A \gg 1$; N_c is a CB density of states and E_D is the depth of the donor level below the CB edge. When muonium is the only shallow-donor centre present, at least within a certain volume, we must put $N_D = 1$. Within this volume, we may safely set $N_A = 0$: since the Mu^0 state is seen, the electron has not dropped to any acceptor level on the μSR timescale. The proportion of dissociated donors, i.e. the Mu^+ fraction, is then $x = n/N_D$, given by

$$\frac{x^2}{1 - x} = \frac{N_c}{2} \exp\left(-\frac{E_D}{kT}\right). \tag{35}$$

With $N_c \propto T^{3/2}$ this reduces to equation (7), as used by Gil *et al* (2001). Their fits typically give $N_c \sim 10^6$, which Weidinger (2001) argues is the number of CB states in a volume bounded by other shallow impurities. For small x, equation (35) approximates to

$$x \propto \exp\left(-\frac{E_D}{2kT}\right), \tag{36}$$

giving about half the activation energy of equation (6), for a given donor depth. In effect, equations (7) and (34)–(36) express loss of electrons to a notional Fermi sea, with the Fermi energy pinned, at the lowest temperatures at least, halfway between the donor level and the CB edge. The establishment of a local Fermi energy by a single muonium centre is clearly open to question, however. Certainly there is no fast equilibration with other shallow donors or spin exchange would broaden the hyperfine satellites beyond detection. It is also noteworthy that the hyperfine satellites remain resolved for the majority of the ionization regime, so neither are there multiple cycles of electron capture and loss, i.e. the dynamic equilibrium between the Mu^0 and Mu^+ states which is a prerequisite of equilibrium expressions such as equation (34). Pending the results of further studies, e.g. on lightly doped material, the inclination of the

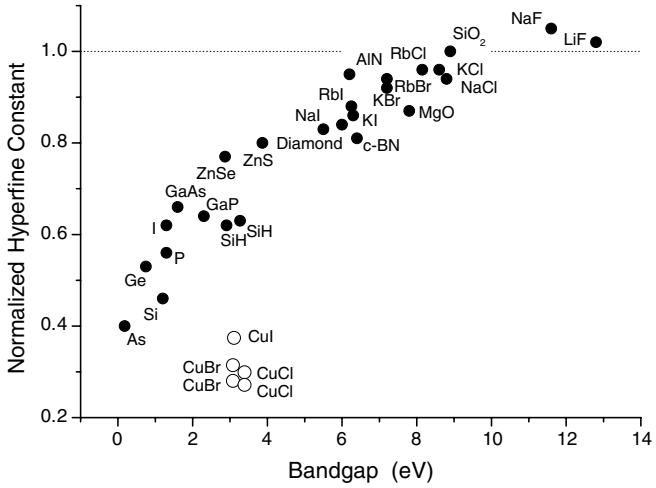

Figure D.1. Hyperfine constants for isotropic muonium centres in semiconductors and dielectrics, relative to the vacuum-state value of 4463 MHz.

present author is to identify the direct ionization energy with the donor depth, according to equation (6).

As a rather different interpretation of ionization regime, it could be that the electron capture—or its cascade down to the observed final state—is somewhat delayed for the shallow states. *Increase* of this delay time with increasing temperature, to values approaching several microseconds, would then account for the reduction of amplitude of the satellites (as polarization is no longer coherently transferred from the precursor Mu^+ to the product Mu^0 signals); it would also account for substantial temperature-dependent phase shifts in the corresponding time-domain oscillations which have been noted in recent data (see footnote 6). This alternative model will be pursued elsewhere.

Appendix D. Compendium of muonium hyperfine constants

Compiled from literature data[19], this is conveniently shown graphically in figure D.1. A broad correlation with bandgap appears for the trapped-atom states, with the notable exception of the cuprous halides (Cox 1987, 2003, Cox and Symons 1986). The deep- and shallow-donor states show no such correlation—contrast figure 22.

References

Abragam A and Bleaney B 1970 *Electron Paramagnetic Resonance of Transition Ions* (Oxford: Clarendon)
Adler D and Yoffa E J 1976 *Phys. Rev. Lett.* **36** 1197–200
Alberto H V, Vilão R C, Piroto Duarte J, Gil J M, Ayres de Campos N, Lichti R L, Davis E A, Cottrell S P and Cox S F J 2001 *Hyperfine Interact.* **136/137** 471–7
Altarelli M and Hsu W Y 1979 *Phys. Rev. Lett.* **43** 1346–49
Atherton N M 1993 *Priciples of Electron Spin Resonance* (Chichester: Ellis Horwood)
Bech Nielsen B, Johannessen P, Stallinga P, Bonde Nielsen K and Byberg J R 1997 *Phys. Rev. Lett.* **79** 1507

[19] With the exception of new values for SiH from R L Lichti and W Nussbaum, private communication. Temperature dependences are not always reported but may be significant. For instance, the hyperfine constant for hexagonal boron nitride varies from 85% of the free atom value at cryogenic temperatures to around 94% at ambient temperature (new data from P J C King, unpublished, not shown in figure D.1).

Bech Nielsen B *et al* 1994 *Defects in Semiconductors II* (*Mat. Res. For.* vol 143–147) ed H Heinrich and W Jantsch (Aedermannsdorf: Trans Tech) p 909
Bech Nielsen B and Grimmeiss H G 1989 *Phys. Rev.* B **40** 12403
Bonde Nielsen K, Bech Nielsen B, Hansen J, Andersen E and Andersen J U 1999 *Phys. Rev.* B **60** 1716–28
Bonde Nielsen K, Dobaczewski L, Søgård S and Bech Nielsen B 2001 *Physica* B **308–310** 134–8
Block D, Herveé A and Cox R T 1982 *Phys. Rev.* B **25** 6049
Breit G 1929 *Phys. Rev.* **34** 553
Brewer J H, Crowe K M, Gygax F M, Johnson R F, Patterson B D, Fleming D G and Schenck A 1973 *Phys. Rev. Lett.* **31** 143–6
Brewer J D, Brewer J H, Morris G D, Eshchenko D G and Storchak V G 2000 *Physica* B **289/290** 428–30
Brewer J H, Crowe K M, Gygax F N and Schenck A 1975 *Muon Physics* vol 3, ed V W Hughes and C S Wu (New York: Academic) pp 3–139
Catlow C R A, Baram P S, Parker S C, Purton J and Wright K V 1995 *Phil. Trans. R. Soc.* **350** 265–76; reprinted in Davis E A and Cox S F J (ed) 1996 *Protons and Muons in Materials Science* (London: Taylor and Francis)
Chow K H 2003 *Physica* B **326** 145–50
Chow K H, Kiefl R F, MacFarlane W A, Schneider J W, Cooke D W, Leon M, Paciotti M, Estle T L, Hitti B, Lichti R L, Cox S F J, Schwab C, Davis E A, Morrobel-Sosa A and Zavieh L 1995 *Phys. Rev.* B **51** 14762
Chow K H, Cox S F J, Davis E A, Dunsiger S R, Estle T L, Hitti B, Kiefl R F and Lichti R L 1997 *Hyperfine Interact.* **105** 309–14
Chow K H, Hitti B and Kiefl R F 1998 *Identification of Defects in Semiconductors* (*Semiconductors and Semimetals*) vol 51A, ed M Stavola (New York: Academic) p 137
Chow K H, Kiefl R F, Hitti B, Estle T L and Lichti R L 2000 *Phys. Rev.* B **84** 2251
Claxton T A and Cox S F J 1994 *Hyperfine Interact.* **87** 917–22 see also Claxton T A, Graham M, Cox S F J, Maric Dj M, Meier P F and Vogel S 1990 *Hyperfine Interact.* **65** 913–26
Cooke D W, Leon M, Paciotti M A, Meier P F, Cox S F J, Davis E A, Estle T L, Hitti B, Lichti R L, Boekema C, Lam J, Morrobel-Sosa A and Dostens J 1994 *Hyperfine Interact.* **86** 699
Cottrell S P, Johnson C, Cox S F J and Scott C A 2003 *Physica* B **326** 248–51
Cox S F J 1987 *J. Phys. C: Solid State Phys.* **20** 3187–319
Cox S F J 1995 *Phil. Trans. R. Soc.* A **350** 227–36
Cox S F J 2003a *Proc. 22nd Int. Conf. on Defects in Semiconductors, (Aarhus, 2003) Physica* B at press
Cox S F J 2003b *Physica* B **326** 113–9
Cox S F J and Symons M C R 1986 *Chem. Phys. Lett.* **126** 516–25
Cox S F J, King P J C, Williams W G, Chow K H, Jestadt Th, Hayes W, Lichti R L, Schwab C R and Davis E A 2000 *Physica* B **289/290** 538–41
Cox S F J, Davis E A, Cottrell S P, King P J C, Lord J S, Gil J M, Alberto H V, Vilão R C, Piroto Duarte J, Ayres de Campos N, Weidinger A, Lichti R L and Irvine S J C 2001a *Phys. Rev. Lett.* **86** 2601–4
Cox S F J, Davis E A, King P J C, Gil J M, Alberto H V, Vilão R C, Piroto Duarte J, Ayres de Campos N and Lichti R L 2001b *J. Phys.: Condens. Matter* **13** 9000–10
Cox S F J, Cottrell S P, King P J C, Lord J S, Alberto H V, Ayres de Campos N, Gil J M, Piroto Duarte J, Vilão R C, Davis E A, Ellis D, Lamb D, Lichti R L, Celebi G and Weidinger A 2002a *Physics of Semiconductors 2002* (*IOP Conf. Series* vol 171) (Bristol: Institute of Physics Publishing) pp 69–75
Cox S F J, Davis E A and Lichti R L 2002b *J. Phys. D: Appl. Phys.* **35** 586–90 plus references therein
Davis E A 1996 The Mott lecture *J. Non-Cryst. Solids* **198–200** 1–10
Davis E A 2001 private communication
Davis E A, Singh A and Cox S F J 1995 *Phil. Trans. R. Soc.* A **350** 227–36
Davis E A, Cox S F J, Lichti R L and Van de Walle 2003 *Appl. Phys. Lett.* **82** 592–4
Dawson W K, Nishiyama K, Macrae R and Nagamine K 1997 *Hyperfine Interact.* **106** 97–103
Eshchenko D G, Storcak V G, Lichti R L and Brewer J H 2003a *Physica* B **326** 120–3
Eshchenko D G, Storchak V G, Cottrell S P and Cox S F J 2003b *Phys. Rev.* B **68** 073201
Estle T L and Vanderwater D A 1983 *Phys. Rev.* B **27** 3962
Estreicher S K 1994 *Mater. Sci. Forum* **148–149** 349–92
Fall C J, Binggeli N and Baldereschi A 2002 *Phys. Rev. Lett.* **88** 156802
Gil J M, Alberto H V, Vilão R C, Piroto Duarte J, Mendes P J, Ferreira L P, Ayres de Campos N, Weidinger A, Krauser J, Niedermayer Ch and Cox S F J 1999 *Phys. Rev. Lett.* **83** 5294–7
Gil J M, Alberto H V, Vilão R C, Piroto Duarte J, Ayres de Campos N, Weidinger A, Krauser J, Davis E A, Cottrell S P and Cox S F J 2001 *Phys. Rev.* B **64** 075205
Gorelkinskii Yu V and Neninnyi N N 1987 *Sov. Tech. Phys. Lett.* **13** 45–6
Gorelkinskii Yu V and Nevinnyi N N 1991 *Physica* B **170** 155–67

Herbert D C and Inkson J 1977 *J. Phys. C: Solid State Phys.* **10** L695–8
Herring C *et al* 2001 *Phys. Rev.* B **64** 125209
Hintermann A, Meier P F and Patterson B D 1980 *Am. J. Phys.* **48** 956–61
Hitti B 1999 *Phys. Rev.* B **59** 4918
Hofmann D M, Hofstaetter A, Leiter F, Zhou H, Henecker F, Meyer B K, Orlinskii S B, Schmidt J and Baranov P G 2002 *Phys. Rev. Lett.* **88** 045504
Hourahine B 2000 *Thesis* University of Exeter
Hugosson H W *et al* 2002 *Phys. Rev.* B **66** 174111
Hutson A R 1957 *Phys. Rev.* B **108** 222
Jones R, Coomer B F, Goss J P, Horahine B and Resende A 2000 *Solid State Phenom.* **71** 173–248
Kadono R 1990 *Hyperfine Interact.* **64** 615–34
see also Kadono R *et al* 1990 *Hyperfine Interact.* **64** 635–40
Kiefl R F, Celio M, Estle T L, Luke G M, Kreitzman S R, Brewer J H, Noakes D R, Ansaldo E J and Nishiyama K 1987 *Phys. Rev. Lett.* **58** 1780
Kiefl R F, Celio M, Estle T L, Kreitzman S R, Luke G M, Riseman T M and Ansaldo E J 1988 *Phys. Rev. Lett.* **60** 224
Kiefl R F and Estle T L 1992 *Hydrogen in Semiconductors* ed S J Pearton, J W Corbett and M Stavola (Berlin: Springer)
Kiefl R F and Kreitzman S R 1992 *Perspectives of Meson Science* ed T Yamazaki, K Nakai and K Nagamine (Amsterdam: North Holland)
Kiliç Ç and Zunger A 2002 *Appl. Phys. Lett.* **81** 73
Krauser J, Riedle Th, Klenk R, Klaer J, Lux-Steiner M C and Weidinger A 2000 *Appl. Phys.* A **70** 617–23
Kreitzman S R *et al* 1995 *Phys. Rev.* B **51** 13117
Kreitzman S R and Roduner E 1995 *Chem. Phys.* **192** 189–203
Lavrov E V, Weber J, Börrnert F, Van de walle C G and Helbig R 2002 *Phys. Rev.* B **66** 165205
Lichti R L 1999 *Hydrogen in Semiconductors II* ed N Nickel (New York: Academic) p 311
Lichti R L 1995 *Phil. Trans. R. Soc.* A **350** 322–33
Lichti R L 2003a *Hydrogen in Materials and Vacuum Systems* (*American Institute of Physics Conf. Proc.* vol 671) ed G R Myneni and S Chattopdhyay (New York: American Institute of Physics) p 55
Lichti R L 2003b *Physica* B **326** 139–44
Lichti R L *et al* 1999 *Phys. Rev.* B **60** 1734
Lichti R L, Cox S F J, Schwab C, Estle T L, Hitti B and Chow K H 1997 *Hyperfine Interact.* **105** 333–8
Lichti R L, Cox S F J, Davis E A, Hitti B and Sjue S K L 2001 *Physica* B **308–310** 73–6
Limpijumnong S and Van de Walle C G 2001 *Phys. Status Solidi* b **228** 303
Look D C, Lu H, Schaff W J, Jasinski J and Liliental-Weber Z 2002 *Appl. Phys. Lett.* **80** 258–60
Lord J S 2003 *Physica* B **326** 129–32
Lord J S *et al* 2001 *Physica* B **308–310** 920–3
Madelung O (ed) 1996 *Semiconductors—Basic Data* (Berlin: Springer)
Mamedov T N *et al* 2000 *Physica* B **289/290** 574–7
Memming M 1983 *Comprehensive Treatise in Electrochemistry* vol 7 (New York: Plenum) p 543
Morenzoni E *et al* 2003 *Physica* B **326** 196–204
Morton J R and Preston K F 1978 *J. Magn. Res.* **30** 577–82
Neugebauer J and Van de Walle C G 1995 *Phys. Rev. Lett.* **75** 4452
Pake G E and T L Estle 1973 *The Physical Principles of Electron Paramagnetic Resonance* (Reading, MA: Benjamin)
Park C H and Chadi D J 1999 *J. Korean Phys. Soc.* **35** S1133
Park C H and Chadi D J 2000 *Phys. Rev. Lett.* **84** 4717
Patel *et al* 1981 *J. Phys. Chem.* **14** 1399
Patterson B D 1988 *Rev. Mod. Phys.* **60** 69–159
Peacock P W and Roberston J 2003 *Appl. Phys. Lett.* **83** 2025
Percival P W, Addison-Jones P W B, Brodovitch J-C, Ji F, Horoyski P J, Thewalt M L W and Anthony T R 1995 *Chem. Phys. Lett.* **245** 90–4
Pratt F L 1997 *Phil. Mag. Lett.* **75** 371–9
Ragnarsson L A *et al* 2001 *Appl. Phys. Lett.* **78** 4169
Schefzik M, Scheuermann R, Schimmele L, Seeger A, Herlach D, Kormann O, Major J and Röck A 2000a *Physica* B **289/290** 511–5
Schefzik M, Schimmele L, Seeger A, Herlach D, Kormann O, Major J and Röck A 2000b *Physica* B **289/290** 521–4
Seager C H, Anderson R A and Estreicher S K 1995 *Phys. Rev. Lett.* **74** 4565
Senba M 1990 *J. Phys. B: At. Mol. Opt. Phys.* **23** 1545–62
Senba M 2001 *J. Phys. B: At. Mol. Opt. Phys.* **34** 4437–54

Shimomura K, Nishiyama K and Kadono R 2002 *Phys. Rev. Lett.* **89** 255505
Shluger A L, Foster A S, Gavartin J L and Sushko P S 2003 Defects in widegap oxides: computer modelling and challenges *Nano and Giga Challenges in Microelectronics* ed J Greer, A Korkin and J Labanowski (Amsterdam: Elsevier) pp 157–222
Schmickler W and Schultze J W 1986 *Modern Aspects of Electrochemistry* vol 17 (New York: Plenum) p 357–410
Simmonds P E, Venghaus H, Sooryakamur R and Dean P J 1982 *Solid State Commun.* **43** 311–4
Slichter C P 1978 *Principles of Magnetic Resonance* (Berlin: Springer) (see also 3rd edn, 1990)
Spaeth J-M 1986 *Hyperfine Interact.* **32** 641–58
Spencer D P, Fleming D G and Brewer J H 1983 *Hyperfine Interact.* **17–19** 567–74
Stoneham A M 1975 *Theory of Defects in Solids* (Oxford: Clarendon) (reprinted as an Oxford Classic, 2001)
Stoneham A M 2001 *Comments at the Institute of Physics Mtg: Magnetic Resonance of Defects in Solids (London, April 2001)*
Stoneham A M and Ramos M M D 1993 *J. Solid State Chem.* **106** 2–12
Storchak V G, Cox S F J, Cottrell S P, Brewer J H, Morris G D, Arseneau D J and Hitti B 1997 *Phys. Rev. Lett.* **78** 2385–8
Street R A 1991 *Hydrogenated Amorphous Silicon* (Cambridge: Cambridge University Press)
Symons M C R 1995 *Protons and Muons in Materials Science* ed E A Davis and S F J Cox (London: Taylor and Francis)
Thomas D G and Lander J J 1956 *J. Chem. Phys.* **25** 1136–42
Van de Walle C G 2000 *Phys. Rev. Lett.* **85** 1012
Van de Walle C G 2001 *Widegap–2001 (Exeter)* unpublished oral presentation
Van de Walle C G 2002 *CMMP–2002 (Brighton)* unpublished oral presentation
Van de Walle C G 2003 *Phys. Status Solidi* b **235** 89–95
Van de Walle C G, Bar-Yam Y and Pantelides S T 1988 *Phys. Rev. Lett.* **60** 2761–4
Van de Walle C G and Neugebauer J 2003 *Nature* **423** 626–8
Walker D C 1983 *Muon and Muonium Chemistry* (Cambridge: Cambridge University Press)
Watkins G D 1981 *Defects in Semiconductors* (*Materials Research Society Symp. Proc.* vol 2) ed J Naryan and T Y Tan (Pitsburg, PA: Materials Research Society) pp 21–30
Wei S-H and Zunger A 1998 *Appl. Phys. Lett.* **72** 2011
Weidinger A 2001 private communication
Wu J, Walukiewicz W, Yu K M, Ager J W III, Haller E E, Lu H, Schaff W J, Saito Y and Nanishi Y 2002 *Appl. Phys. Lett.* **80** 3967
Yu E T, McCaldin J O and McGill T C 1992 Band offsets in semiconductor heterojunctions (review) *Solid State Physics* vol 46, ed H Ehrenrich and D Turnbull p 1
Yu P Y and Cardona M 1996 *Fundamentals of Semiconductors* (Berlin: Springer)
Zunger A 1986 *Solid State Physics* vol 39, ed F Seitz, D Turnbull and H Ehrenreich (New York: Academic) p 275