both Si and O. The band limits have been analysed in terms of TB interactions by Laughlin *et al.* [327]. There is little *s–p* hybridization at the oxygen site and the O 3*s* states lie at −22 eV. There are three types of O 2*p* states. The O $p\pi$ states are normal to the S–O–Si bonding plane and form the upper valence band. The O *p* orbital along the Si–Si direction bonds strongly to the Si sp^3 orbitals and produces the deep bonding states from −5 to −10 eV. The remaining *p* orbital interacts weakly with the Si sp^3 hybrids and produces states from 0 to −5 eV, overlapping the true π states [323]. The valence maximum is not at Γ, producing an indirect gap [326].

The nature of the conduction band is a matter of some contention. The conduction band minimum has a low effective mass and contains Si 3*s* states [326]. The Si *p* states form the higher conduction bands. In common with l-Ar or CH_4, the mobility edge of SiO_2 is very close to the conduction minimum [334]. The high electron mobility in l-Ar is attributed to the *s*-like character of its conduction band minimum. Such isotropic states are sensitive only to bond length disorder, which is small in l-Ar. In a-SiO_2, Mott emphasizes the contribution of O 3*s* states to the conduction minimum which would also be insensitive to disorder because of the rigidity of the Si tetrahedral angle [326, 334]. However, this is a simplification of the real situation as Si 3*s* states are the dominant contribution. Ching [319, 330] finds very delocalized states at the minimum in a self-consistent LCAO calculation of a-SiO_2, implying that a high degree of cancellation must occur within the effective potential seen by the Si 3*s* states themselves.

The importance of Si 3*d* states in SiO_2 has also been contentious. The Si 3*d* orbital is an excited state whose extent depends strongly on the bonding potential. Early LCAO calculations on SiO_4^{4-} clusters sometimes find significant $pd\pi$ hybridization and *d* content in the $p\pi$ band but bulk pseudopotential results find *d* effects to be less important [326].

A final source of contention is the nature of the gap in SiO_2. The minimum gap in quartz is now believed to be about 9 eV, the onset of photoconductivity [334, 335]. The gap is indirect as the valence maximum lies at the zone boundary, *M*. Also, transitions for the higher π valence bands are forbidden in the dipole approximation in any form of SiO_2 with tetrahedral Si sites because the phases of these π orbitals prevent their coupling to the Si *s* conduction states [336]. Thus band-to-band optical transitions below 12 eV are very weak in SiO_2. The optical spectrum of all forms of SiO_2 exhibits a large peak at 10·2 eV which can now be assigned to a direct exciton. Its binding energy is difficult to determine because of the complexities of the band-to-band transitions, but values in the range 1·5 eV are likely [334].

a-SiO_x alloys are very interesting because of the large difference in band gaps of SiO_2 and Si [337–339]. The band edges of Si lie entirely within the gap of SiO_2. The σ and σ^* states of isolated Si–Si bonds also lie within the gap of SiO_2 [340]. It is then possible to study the localized Si–Si states within the insulating SiO_2 matrix.

GeO_2 has two polytypes. Hexagonal GeO_2 has a 4 : 2 co-ordination similar to SiO_2 but the O angle has decreased from 144 to 130°. The second polytype has the tetragonal (rutile) structure of SnO_2 in which Ge is octahedrally co-ordinated and O has a planar trigonal co-ordination. This configuration is more ionic than the hexagonal form. Interestingly, the valence band maximum is still formed of oxygen π states because its site is planar [341]. a-GeO_2 has a 4 : 2 co-ordination as in the hexagonal form. The bond angle distortion at the O site has decreased from ±65° in SiO_2 to ±37° in GeO_2 [342]. The decreased angle and distortions can both be attributed to the greater Ge radius.

The highly important silicate glasses are formed by fusing and quenching SiO_2 with various alkali and alkaline earth oxides. Only a brief discussion of their properties is included here. The effect of the oxides is to break up the network by creating non-bridging oxygen sites, which are monovalent and negatively charged. The precise configuration of the metal ions has been the subject of many diffraction and Raman studies. [343, 344], but only recently has detailed information from EXAFS become available [345]. EXAFS finds that the metal or 'modifier' ions do not enter random, interstitial sites in the network but have quite precise configurations, adjacent to the non-bridging oxygens and similar to those in the analogous crystalline silicates [345]. The effect of modifiers is to reduce the optical gap [346]. The non-bridging oxygen sites broaden the valence band upwards by 1·5–2 eV in XPS due to their negative charge [340, 347]. Any further decrease in gap can be attributed to modifier states lowering the conduction band edge [348]. Modifier oxides have also been added to a-GeO_2. The metal ions occupy similarly precise configurations. Interestingly, modifiers such as Li_2 or Rb_2O can induce the Ge site to change to the more ionic sixfold configuration [349].

7.4. Si_3N_4

Silicon nitride is an insulator which possesses a valence maximum of π states purely as a result of geometrical restrictions [350]. The size mismatch of Si and N atoms produces Si–Si repulsions around the N site and cause it to be planar. The N π orbital normal to this plane acts as a lone pair, even though N has p^3 rather than p^4 configuration. The calculated DOS [350, 351] disagrees with the only XPS data available [352]. Silicon nitride is prepared by vapour decomposition of SiH_4 and NH_3 and contains chemically bonded hydrogen, like a-Si : H. This acts in a passivating role as in a-Si : H.

7.5. *Amorphous III–V semiconductors*

Crystalline III–V semiconductors have played a major role in our understanding of the tetrahedrally bonded semiconductors. The rdf shows that amorphous III–V semiconductors are also tetrahedrally bonded [353]. However, they have attracted comparatively less interest for various reasons: the influences of various forms of disorder is not fully settled, substitutional doping has not been observed and perhaps also because the chalcogenides have acted as the standard compound semiconductor. A major question in amorphous III–V semiconductors is whether the major form of bonding disorder is due to dangling bonds or homopolar 'wrong' bonds. Experimentally, amorphous III–V semiconductors can be prepared by flash evaporation, sputtering or in hydrogenated forms [353–357], as with a-Si : H. Clearly, stoichiometry control is crucial when studying the proportion of wrong bonds.

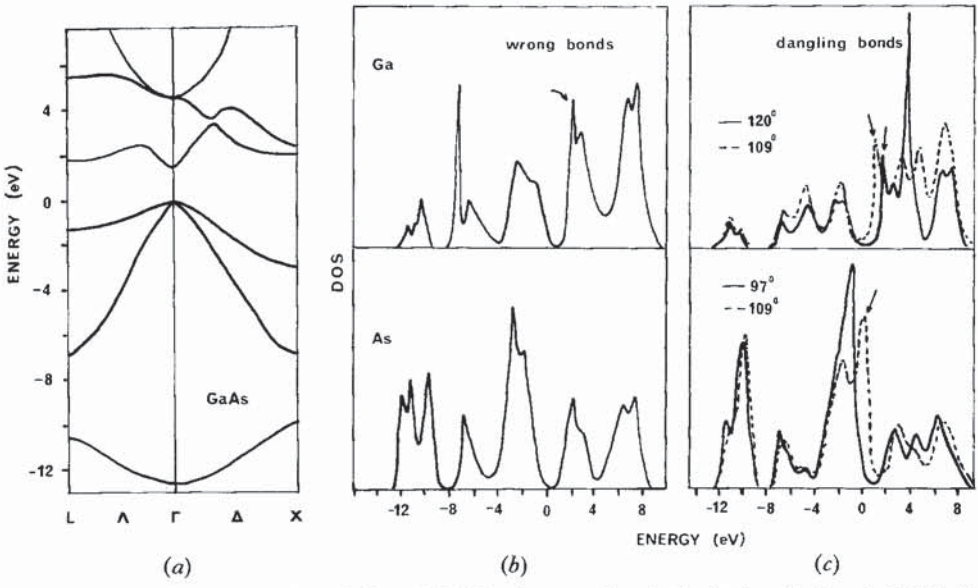
As in Si, the tetrahedral bonding of a-GaAs is due to sp^3 hybrids but Ga and As are naturally trivalent, unlike Si. Tetrahedral bonding requires a formal charge transfer to give Ga^-As^+ so that each site now possesses 4 electrons, one per sp^3 hybrid [76]. The ions Ga^- and As^+ are now tetravalent in the 8-N rule. Further charge transfer occurs during bonding which produces slightly negative As sites, but formally GaAs is written Ga^-As^+. This electron counting requirement must be remembered when considering possible forms of disorder.

The bands [54] of c-GaAs are shown in fig. 32 for reference. GaAs is not strongly polar and the states generally show mixed Ga and As character. The lowest valence band is the exception as it is strongly As *s*-like. Ga *s* and As *p* states interact in the second valence band region while Ga *p* and As *p* states interact in the upper valence region. The conduction band minimum of c-GaAs is at Γ, giving a direct gap, and overall the lowest conduction bands are mostly Ga *s*-like. The structure of chemically ordered a-GaAs is similar, but with a smearing of conduction band features (see, e.g. GaAs in the BC-8 structure [358]).

We discuss wrong and dangling bonds in two stages, first the experimental situation is covered here and then a qualitative theory is developed in §8.9 using analogy to surface reconstructions. The a-Si network is known to contain odd-membered rings from its *s* band DOS, and such a network for a-GaAs must contain homopolar bonds. A dangling bond is the energetically least favourable defect in a-Si. The antisite defect with wrong bonds is energetically more favourable than the vacancy in c-GaAs. Following similar arguments, wrong bonds were expected to out-number dangling bonds in a-GaAs and evidence for wrong bonds was sought in the rdfs, XPS, UPS, optical spectra and EXAFS [353–366]. Only Connell noted that relaxed dangling bonds (three-fold sites) were to be expected at void surfaces [364].

Connell and Temkin [33] constructed a four-fold network containing only even-membered rings, but then Temkin [29] showed that very careful diffraction measurements are needed to prove the necessity for such a network from the crucial 4·5 Å region of the rdf of a-GaAs. Homopolar bonds are more easily detected in the skew III–V semiconductors (because they differ in length). Shevchik [354] found a wider first peak in the rdfs of a-GaP, a-GaSb and a-InSb, suggesting wrong bonds.

Fig. 32



(*a*) (*b*) (*c*)

(*a*) Bandstructure of c-GaAs. (*b*) Local DOS of wrong bonds in GaAs. (*c*) Local DOS of dangling bonds in GaAs; unrelaxed for bond angles of 109° and relaxed angles of 120° (Ga) and 97° (As). Gap states are arrowed.

EXAFS is also much more sensitive to wrong bonds in skew III–V semiconductors because the larger differences in both lengths and phase shifts help the analysis. Wrong bonds have been detected in a-GaP but possibly not in a-GaAs [361]. Core level energies vary with a site's associated valence charge and are thus sensitive to the introduction of homopolar bonds. Shevchik *et al.* [359] found that the core levels of many III V semiconductors had a similar energy and width in the crystal and amorphous phases. This suggests little 'wrong' bonding unless some subtle screening process occurs.

We suggest that threefold sites rather than wrong bonds are the predominant defect in amorphous III–V semiconductors. The difference between Si and GaAs is illustrated by the behaviour of their surfaces; most Si surfaces change their periodicity (reconstruct) to reform bonds, while GaAs(110) only relaxes (§8.9). Thus, the analogy between dangling bonds in a-Si and a-GaAs is false. Si is naturally tetravalent, while Ga and As are trivalent. Broken bonds in a-Si usually reform but 'dangling bond' sites in a-GaAs now have their natural valence and can relax towards their bond angles in a-Ga or a-As. This relaxation stabilizes threefold sites. Such relaxation at similar sites at vacancies in c-GaAs is opposed by the stronger bond stretching forces and consequently antisites are the more stable crystalline defect. This would suggest that tetrahedral Ga and As sites are only stable in a-GaAs when surrounded by heteropolar bonds.

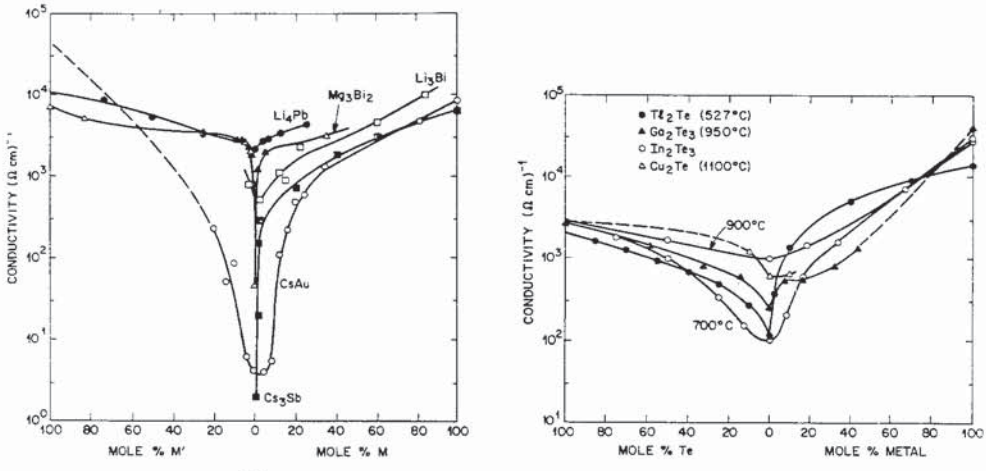
The local DOS at wrong bond and dangling bond sites has been calculated using the recursion method [365] and sp^3s^* TB interactions [99], (table 3, fig. 32 (*b*)). In GaAs, the As–As bond does not produce gap states, as its σ and σ^* states lie on either side of the gap. It does produce a characteristic twin-peaked structure in the As *s* band at -12 eV. The Ga–Ga bond does produce gap states. Both Ga and As dangling bond sites give rise to gap states when they possess unrelaxed 109·5° bond angles. Upon relaxation to an 87° angle the As state moves downwards into the valence band, and the Ga state moves towards the conduction band edge as it relaxes to 120°, but remains just in the gap.

The behaviour of amorphous III–V semiconductors away from stoichiometry depends on the stability of wrong bonds. If an excess enters substitutionally, it forms an antisite with wrong bonds. Antisites produce deep gap states in most III–V semiconductors so little self-doping occurs. However, a more likely event is that the excess enters trivalent sites: These are self-compensated [246] and leave ε_f in the gap centre, as seen experimentally in a-GaSb [355]. Finally, we note that the optical spectra of III–V semiconductors develops into a single red-shifted peak on becoming amorphous [355, 366], as in Si.

7.6. *III–VI systems*

The effects of variable valence in amorphous and liquid semiconductors are most clearly seen in group III elements and their compounds. There is a steady change from trivalence to monovalence in the tellurides with increased atomic number of the group III element. The monovalent tellurides are ionic while the trivalent tellurides are covalent. Ordering only occurs at $x = 0{\cdot}67$ in l-Tl_xTe_{1-x} (fig. 33 (*b*)) [367, 368]. The Te sites have co-ordination of about 9 in l-Tl_2Te, implying a structure similar to disordered antifluorite [369]. The rdf is almost independent of temperature [231], as expected for an ionic phase. In contrast, l-Ga_xTe_{1-x} shows its most pronounced ordering at $x = 0{\cdot}4$ [370, 371]. It is semiconducting at $x = 0{\cdot}4$, just above the melting

Fig. 33



(*a*) (*b*)

Composition dependence of the conductivities of (*a*) some ionically bonded liquid semiconductors and (*b*) some liquid tellurides. The alloys are presented using normalized compositions by expressing them as alloys of the pure elements and the stoichiometric compound, as in [10].

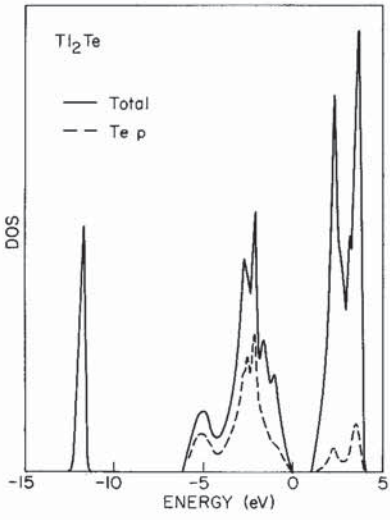
point, and its composition is consistent with covalent bonding. In contrast to l-Tl_2Te, the rdf of l-Ga_2Te_3 is strongly temperature-dependent [372]. This system also undergoes an MNMT with increasing temperature which can be followed by resistivity, thermopower and NMR experiments [370–372]. The most interesting interpretation of the transition is that it results from a decreasing stability of covalent trivalent Ga [372]. Its structure transforms to that of l-Tl_2Te plus additional Te sites.

The electronic structure of l-Tl_2Te has recently been calculated for an ionic model (fig. 34) [373]. As the Tl *s* electrons do not ionize, the lowest conduction band is a π *p* state, while the Tl *s* states hybridize with the Te *p* states in the upper valence band. The Tl-rich alloys are metallic, and the excess Tl atoms enter as Tl^+, plus free electrons [368]. The Te-rich alloys are semiconducting at low temperatures so Te must enter covalently. Thus Te-rich alloys must contain Te_n^{2-} chains. The paramagnetism of these alloys suggests that additional neutral chain end sites are present [374, 375]. At elevated temperatures and in very Te-rich alloys threefold Te sites appear giving semimetallic properties.

The simplest semiconducting model of l-Ga_2Te_3 is an analogy to the defect zincblende structure of c-Ga_2Te_3. In this lattice one-third of the Ga sites are vacant, so Ga always has four-fold co-ordination while Te has mixed two- and threefold co-ordinations. This combination of sites produces an electronic structure very similar to that of GaAs, with a Ga *s*-like lower conduction band [376]. An increase in mean Te co-ordination drives the MNMT, as this causes Ga *s* states to move into the valence band, giving the valence change and also tending to wash out the gap.

l-In_xTe_{1-x} alloys order at both $x=0{\cdot}4$ and $0{\cdot}67$ [370, 377–379]. Again, monovalent In sites are ionic, and trivalent In sites are covalent but with a decreased stability at higher temperatures.

Fig. 34



TB DOS for Tl_2Te in an antifluorite structure.

The a-Tl_xTe_{1-x} alloys exhibit some of the most dramatic ordering seen in amorphous systems at compositions given by $x=2/(n+2)$, for integer n [380]. An interpretation of this ordering in terms of Tl^+ and Te_n^{2-} units has been sought [373].

7.7. *Ionic semiconductors*

Mg_3Bi_2, Cs_3Sb and CsAu are amongst the best known charge-transfer semiconductors. The calculated TB ionicities of the metal compounds of group V elements (pnictides) ranges from $f_i=0{\cdot}05$–$0{\cdot}65$ and all the pnictides possess a composition-dependent MNMT, which appear to follow the band-filling model described in §6.2 [264]. The alkali-gold salts such as CsAu are extremely ionic with $f_i=0{\cdot}97$. CsAu is so ionic that ionic conductivity dominates in l-CsAu [381] and its MNMT occurs in localized defect states rather than in its band tails, producing semiconducting alloys over a wider than usual composition range [264, 382, 383].

l-Mg_xBi_{1-x} and l-Mg_xSb_{1-x} were the first ionically ordered alloys to be studied [384]. They exhibit the classic signatures of ionic bonding such as a strong, sharp dip in the conductivity and a sign reversal in the thermopower as x varies. The amorphous phases are similar, in general [385, 386]. Thus, in the band-filling model excess Mg enters as ions:

$$Mg \rightarrow Mg^{2+} + 2e^-$$

and Bi as

$$Bi \rightarrow Bi^{3-} + 3h^+.$$

The electronic structure of their crystals is as expected for an ionic compound; the lower conduction bands are largely formed from Mg s states and the upper valence

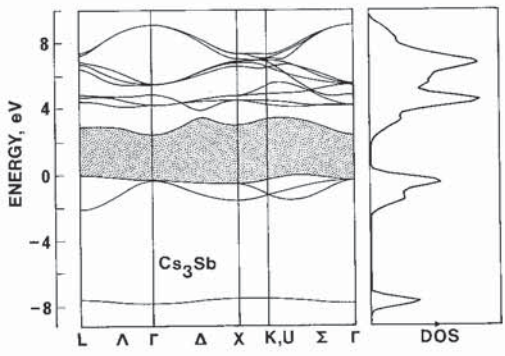
band is due to Bi p states. The calculated TB ionicities are close to 0·60 [264]. The only modification from this simple model occurs in a-Mg_xSb_{1-x} due to the semiconducting nature of a-Sb. Thus while excess Bi enters a-Mg_3Bi_2 as an anion, excess Sb forms covalent bonds within the usual ionic matrix and the optical gap remains for $x < 0{\cdot}4$ [246, 387].

The alkali pnictides crystallise in the cubic Li_3Sb or hexagonal Na_3P structures. The electronic structure of a range of pnictides has been calculated by the TB method for the cubic modifications [264]. Their ionicities are found to depend strongly on the cation but weakly on the anion, such that Cs salts have ionicities of roughly 0·60 but Li salts are less than 0·10. This low ionicity accords with the intuition that an Sb^{3-} should be highly screeend, except in salts of the extremely electropositive Cs. Their electronic structure is that expected for salts with anions with completed octets (fig. 35); the lower conduction bands are formed of metal s states and the upper valence band from anion p states. A notable aspect of the band structure is that an optical gap is found in spite of the presence of many metal–metal nearest neighbours, required by the stoichiometry. A similar electronic structure and SRO is expected for the liquids. The liquids exhibit composition-dependent MNMTs which have been studied experimentally by conductivity, thermopower, magnetic susceptibility and NMR measurements [388, 389]. The conductivities drop sharply on the metal-rich side as stoichiometry is approached indicating that the MNMT occurs in the band tails, in accord with the band-filling model (figs. 26, 33 (*a*)). Direct evidence for localization in band tails is seen by NMR [388].

A number of liquid alloys and evaporated films of alkalis and group IV elements also show strong ordering, reminiscent of ionic bonding [390–397]. Low conductivities are found in some of the films. Neutron diffraction shows that strong ordering occurs in l-Li_4Pb [397] and in a statistical model of its structure [391]. However, its conductivity remains in the metallic regime and the charge transfer is expected to be small [264].

The alkali-gold salts are perhaps the most remarkable ionic compounds known. They crystallize in the CsCl or B2 structure. As both constituents are univalent, their ionicity results purely from differences in the pseudopotentials caused by the weakly bound d electrons of the noble metal Au. CsAu has recently become a test-bed for

Fig. 35



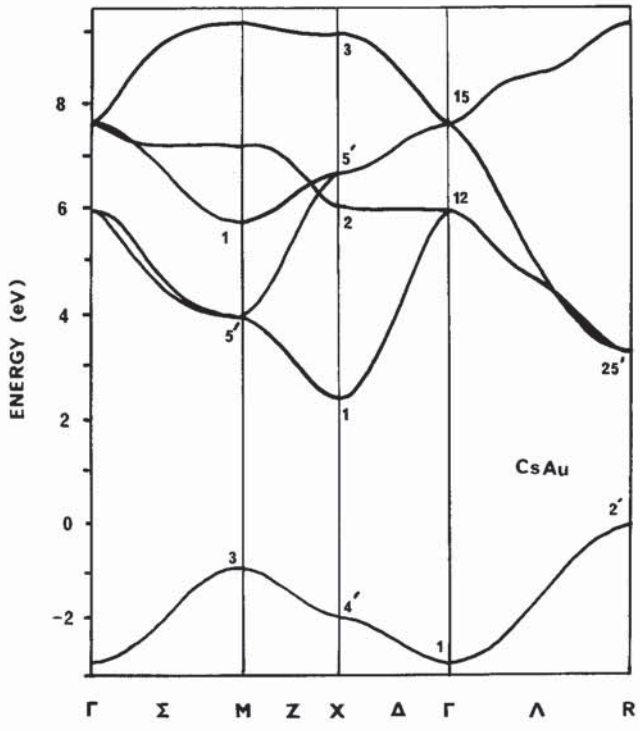
TB bands and DOS for Cs_3Sb (after [264]).

advanced band structure techniques in view of the difficulty of treating the Au 5*d* orbitals and the huge charge redistribution. A consensus of the band structure of CsAu has recently emerged from three calculations using LDF screened atomic pseudopotentials [398–400] and their bands were used as the basis of a TB fit [264].

The fitted TB bands are shown in fig. 36. The single valence band is strongly localized on the Au 6*s* orbitals [398–404], to the extent that the TB ionicities of KAu, RbAu and CsAu exceed 0·95. The lowest conduction band is formed largely of Au 6*p* states (in X_1, Γ_{15} and R'_{25}) followed by the Cs 5*d* states in Γ_{12}. Although a Cs 6*s* to Au 6*s* charge transfer produces the high ionicity, the Cs 6*s* states are repelled above the Cs 5*d* states. The contrast in electronic structures of 2-electron or 'suboctet' CsAu, octet NaCl and 12-electron or 'superoctet' Tl_2Te is very revealing. While ionic systems with 8-electrons per formula unit have valence and conduction bands localized on the anions and cations, respectively, suboctet compounds tend to have both bands localized on the anion while the band-edge states of superoctet compounds are predominantly localized on the cation. A remarkable feature of the ionicities of *s*-bonded alkali-gold salts is that their very high ionicities can exist in compounds with band gaps of 2 eV and less. By comparison, band gaps of over 10 eV are required to produce similar ionicities in octet compounds like CsCl.

l-CsAu has been one of the most extensively studied ionic semiconductors [405–407]. Structurally [408], strong ionic ordered is confirmed by the presence of a prepeak in its total structure factor [249]. The electron localization on the

Fig. 36



TB bands for CsAu parameters in table 6.

semiconducting side of the MNMT is observed directly by NMR [382], In the Cs-rich alloys, NMR is also able to show that the excess electrons are localized between atoms rather than in atom-centred states. This is direct evidence for localization in defect states due to anion vacancies (*F* centres), rather than in band tail states. An electron is trapped at an anion vacancy in a crystalline ionic compound by the Coulombic force which is enhanced by the inward lattice relaxation. *F* centres have been shown to be stable in l-CsAu because of its excellent chemical ordering, large heat of formation and large molar volume contraction compared to its elements [383]. Thus, large ionicity rather than a crystalline lattice is the prerequisite for *F* centres. The idea of *F* centres is supported by the wider range of semiconducting compositions in l-Cs_xAu_{1-x} [264]. If this system followed a band-filling model of § 6.9 only alloys of compositions of roughly $0{\cdot}45 < x < 0{\cdot}55$ would be semiconducting. The wider conductivity minimum seen experimentally (fig. 33) is evidence that a mechanism of localization, additional to 1-electron disorder, contributes in l-CsAu. In the present case the mechanism is the lattice relaxation around the vacancy, which is polaronic.

The defect-driven MNMT of the highly ionic alkali-gold liquids distinguishes them from other liquid semiconductors which have lower ionicities and conventional pseudogap states. F-centres are also believed to exist in other molten salts such as l-CsCl [409]. As in l-CsAu, electrons from excess Cs are found by NMR to be localized in vacancy-like orbitals rather than in band tail states [410]. This invalidates theories of the MNMT in metal–salt systems based on 1-electron approaches [411, 412]. Again, ionicity and strong ordering has been sufficient to stabilize defects generally associated with the crystalline state.

§ 8. Defects and impurities

8.1. *Why defects?*

Spurred by the discovery of doping in a-Si : H, it is now increasingly appreciated how many electronic properties of amorphous semiconductors can be separated into intrinsic and extrinsic parts, as was possible for crystals. Defects in crystals are a more easily accepted concept because the lattice periodicity provides the reference defect-free state. In random networks, bulk co-ordinations and bond lengths act as this reference state.

The bulk co-ordination determines both the density and the electronic behaviour of network defects. If we define defects at sites capable of producing states deep in the pseudogap of an amorphous semiconductor, this separates the major perturbations such as broken bonds or large bond length changes from the other forms of disorder which only produce band-edge fluctuations. The defect density is related to how a network chooses to apportion its strain between these two types of perturbations. Our ideas on this aspect have changed recently [40, 131, 413]. Our early ideas emphasized network flexibility which clearly decreases with increasing co-ordination. As it appeared possible to build four-fold co-ordinated random networks without strain accumulation [34], it was assumed that strain could always be accommodated by variations in angles and ring statistics. As a broken bond is the most energetically costly defect of a random network, earlier workers emphasized the need to minimize the number of broken bonds by network rearrangements and

bond reconstruction, even if only long, weaker bonds resulted [131, 413]. This self-healing property of the network expels states from the gap. In contrast, Phillips considered the number of degrees of freedom of a network obeying valence force-field energies. It was shown that networks with an average co-ordination $N_c < 2{\cdot}4$ were overconstrained and poor glass formers [40, 414–416]. He suggested that the excess strain in overconstrained networks such as a-Si was relieved by broken bonds at internal surfaces. Thus, he proposed that both defects and intermediate-range order were intrinsic to the amorphous state. In chemically-ordered binary alloys like a-$GeSe_2$, the homopolar bonds also acts like a dangling bond in reducing the network strain. The relative importance of homopolar bonds and dangling bonds in binary alloys, and their relationship to intermediate range order is discussed briefly in § 8.9.

8.2. *Defects in chalcogenides and pnictides*

The effect of co-ordination on the electronic behaviour of defects is best described using the idea of effective electronic correlation energy or U, which measures the degree of reconstruction [12–16, 417–422]. U is defined by the Hubbard hamiltonian:

$$H = \sum_{i,\sigma} E_i n_{i\sigma} + U_i n_\sigma n_{-\sigma}, \tag{42}$$

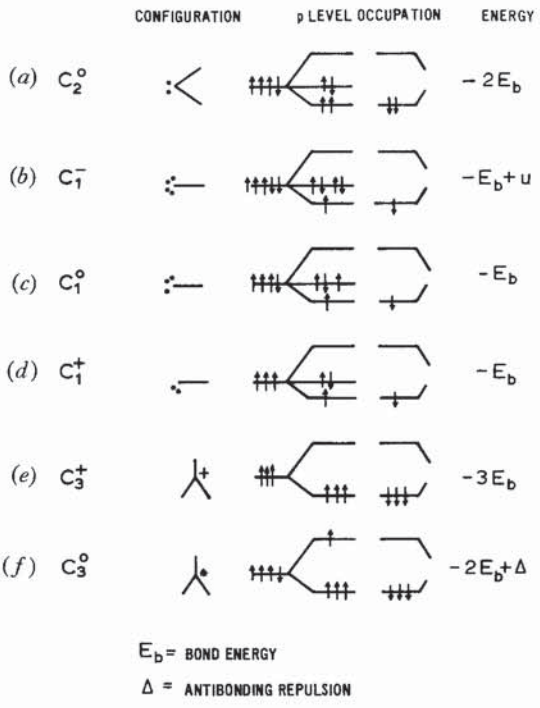
where n_σ is the occupation number of spin σ at site i, E_i are the one electron energies and U_i the on-site correlation energy. A negative U describes an electron attraction. Anderson [417] showed that the pairing-up of electrons into covalent bonds provides an effective electron–electron attraction or negative U which is sufficient to overcome their usual Coulombic repulsion. The formation of covalent bonds is the driving force of many reconstructions. In this manner, a strong electron–lattice interaction translates into an electron–electron attraction.

Consider the situation for dangling bonds. The neutral dangling bond is singly occupied (D^0). The effect of a negative U is that only the diamagnetic states D^+ and D^- of a dangling bond are stable, while the paramagnetic state D^0 is metastable and decays into D^+ and D^-. For $U > 0$ all three stages of the dangling bond, D^+, D^0 and D^- are stable and observed. Thus, the change in sign of U from positive in a-Si to negative in a-Se naturally explains how ESR signals are only seen in a-Si in the cold and dark, while a-Se is always diamagnetic even when containing impurities of different valence [423]. We now show how the lone pairs of Se and As allow the major bond reconstructions which produces their negative U.

The origin of negative-U defects is best appreciated from the bonding models of Mott, Kastner and coworkers [418–421]. The defect site is labelled according to element, C (chalcogen), P (pnictide) and T (tetrahedrally bonded element of group IV), with the charge as superscript and co-ordination as subscript. Thus we have T_4^0, P_3^0 and C_2^0 for normal valence.

The negative U of chalcogen defects arises from the overcoordination of the D^+ site. We recall that the valence configuration of chalcogens (s^2p^4) produces occupied s^2 non-bonding states, σ and π states and unoccupied σ^* conduction states at C_2^0 sites (fig. 37). The π states are in the upper valence states. At the neutral chain end, C_1^0, σ and σ^* states have been converted into an extra π state. Its negative state C_1^- (or D^-) is inert and spin-paired. In contrast, its positive state C_1^+ possesses an unoccupied π orbital which is capable of accepting a dative bond. An adjacent C_2^0 site donates its π

Fig. 37



Configurations and bond energies of defects in chalcogens.

electrons to form the bond. Representing a bond as a dash and the π electrons as two dots, the reaction is

$$-C^+ + :C= \rightarrow -C-C^+= \tag{43}$$

$$C_1^+ + C_2^0 \rightarrow C_2^0 + C_3^+, \tag{44}$$

where C_3^+ is the D^+. U can be defined by the Hubbard hamiltonian (42) or equivalently by heat of the disproportionation reaction:

$$2D^0 \rightarrow D^+ + D^-. \tag{45}$$

This is exothermic for $U<0$. We find $U=-2E_b$ by considering bond energies (fig. 37).

We also see how the reaction sequence:

$$2C_2^0 \rightarrow 2C_1^0 \rightarrow C_1^+ + C_1^- \rightarrow C_3^+ + C_1^- \tag{46}$$

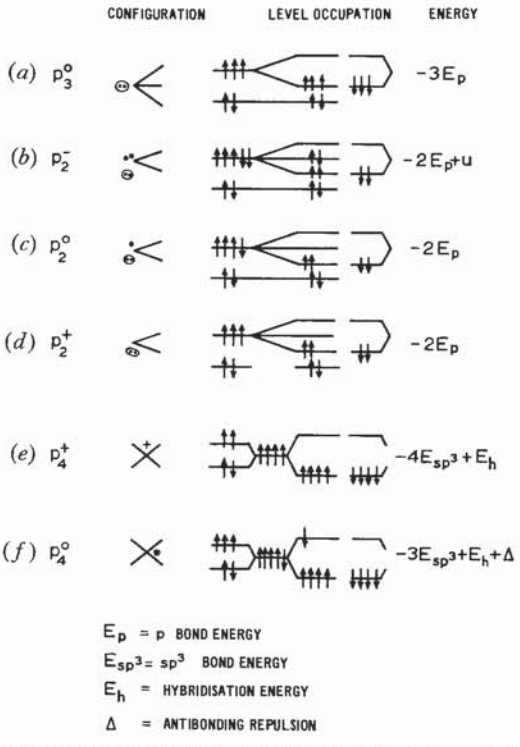
initially breaks a bond and then reforms a bond, a process christened 'bond alternation' [420]. The defects C_3^+ and C_1^- appear in pairs called valence alternation pairs or VAPs. The C_3^+ site is more stable than C_1^+ because of the extra bond. The VAP creation energy is quite low, equal to the true correlation energy U, if the energy of the bonds are taken to be equal, because the number of bonds is the same as in the defect-free state. The creation energy is much smaller than that of simply breaking a bond, and results in a large defect density.

The C_3^+ and C_1^- give deep gap states near the conduction and valence edges, respectively [418]. An electron placed at a C_3^+ site must enter a σ^* state, which is pulled down into the gap from the conduction band by its positive charge. The extra electron at C_1^- is a π electron and the negative charge repels the state above the valence edge. Realistic calculations, including band broadening and Coulombic effects, find that C_3 has a moderately localized state in the upper gap, as expected [424]. The C_1 state is very localized and is found quite far above the valence edge due to an important self-energy shift of its orbital energies. C_1^0 is frequently called an amphoteric centre because it can accept or donate an electron, becoming C_3^+ and C_1^-, respectively. An intermediate step could be written as C_3^0. The presence of a stabilizing extra bond and a destabilizing σ^* electron counterbalance each other in the simple bonding models. Detailed calculations favour the C_1^0 as more stable than C_3^0 [425]. The ESR signal of the D^0 site also shows it to be a simple dangling bond, C_1^0 [440].

The pnictides are also able to form dative bonds and negative-U defects [426]. Atoms must be capable of acting as electron-pair donors to give $U<0$. The role of the π electrons in chalcogens is assumed by the non-bonded s electrons in pnictides. Their participation requires promotion to an sp^3 configuration and the promotion energy reduces the stability of negative U defects in pnictides. Following fig. 38, the overcoordination of D^+ is

$$P_2^+ + P_3^0 \rightarrow P_3^0 + P_4^+ . \tag{47}$$

Fig. 38



Configurations and bond energies of defects in pnictides.

The P_4^+ and P_2^- sites constitute the VAP in As. The absence of any non-bonded valence electrons in group IV elements prohibits their overcoordination within a scheme of *sp* pair bonds, while valence expansion using excited orbitals is not seen in these systems. Thus, Kastner *et al.* [420] explain the qualitative difference of defects in groups IV, V and VI in terms of a requirement for non-bonded electrons.

Defect levels in pnictides differ slightly from those in chalcogens because the valence band is now σ-like rather than π-like [426, 427]. The P_2 dangling bond state is now expected near mid-gap rather than near the valence edge. More detailed calculations verify this and find the defect states to be less localized than in the chalcogens [428–430].

8.3. *Defect processes*

The negative-U defects provide a cogent interpretation of many processes in amorphous chalcogenides and pnictides [431–434], although other theories exist [435]. In amorphous chalcogenides ε_f is firmly pinned near mid-gap indicating a significant density of gap states. These cannot be the usual dangling bond states because there is no evidence of their paramagnetic configuration from magnetic susceptibility or in cold dark ESR measurements. Adding electrons to a defect system of positive U by raising ε_f produces the sequence of occupations D^+, D^0, D^-. The instability of D^0 for $U<0$ leads to the immediate occupation of D^- after D^+. Considering the Hubbard hamiltonian (42), we find that ε_f is pinned at an energy U which is approximately mid-gap because $U=-2E_b$ [417].

The lowest transitions of the D^+, D^- system involve electron pairs. Pinning of ε_f is frequently expressed as a pinning due to a high density of 2-electron states. Ac conductivity in chalcogenides is interpreted as the hopping or tunnelling of paired electrons [436]. Direct experimental evidence for electron pair states or overcoordinated defects is not available for amorphous semiconductors, but an increasing number of defect processes in c-Si are being interpreted as negative U defects which can be studied experimentally in a much more direct manner [437–439].

Photoexcitation of defects produces the clearest evidence for the negative U model [440–442]. Photoabsorption at the band edge in chalcogenides (fig. 39) excites an electron from D^- giving a D^0 site and a free electron. The D^0 sites are visible in ESR, produce an absorption down to mid-gap and cause a luminescence at the mid-gap energy when the electron recombines at D^0. Light-induced ESR gives structural details of the paramagnetic state of the defects. These are found to be chain ends in S and Se, expected. Two types of centres are seen in arsenic chalcogenides, As_2^0 and Si_1^0, by the reactions:

$$-\mathrm{Se}^- \rightarrow -\mathrm{Se}^0 + e^-,$$

$$=\mathrm{Se}^+ - \mathrm{As} = + e^- \rightarrow = \mathrm{Se} + \mathrm{As}^0. \tag{48}$$

A photoinduced and dark ESR signal is seen in a-As itself because defects with both $U>0$ and $U<0$ exist [440].

Photoluminescence has been interpreted as both a single defect process [12, 433, 434] and an internal process in a paired D^+D^- defect or intimate VAP (IVAP)) [443–445]. In the single defect model, largely due to Street, the absorbed photon excites an electron from D^- or the valence band to the conduction band which then thermalizes down its tail. Meantime, the D^- has relaxed to a D^0 configuration and luminescence

Fig. 39

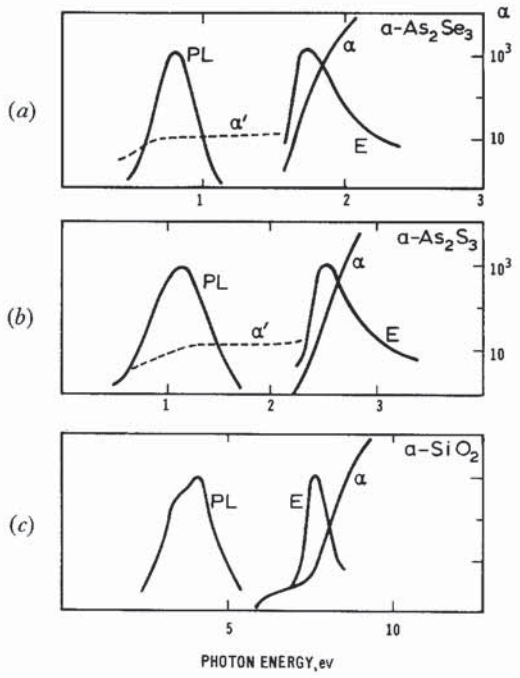


Photo-induced luminescence and absorption in a-As_2Se_3, a-As_2S_3 and a-SiO_2 (after [459]).

occurs by recombination of this electron at D^0. The Stokes shift between excitation and luminescence energies arises from the lattice relaxation in $D^- \rightarrow D^0$. If the electron hops away from the D^0, non-radiative recombination results.

The charges of D^+ and D^- can cause them to pair up into dipolar IVAPs in hosts of low screening. In the IVAP model [445], the C_3^+–C_1 pair (in a-Se) is excited by electron transfer to the self-trapped exciton configuration C_3^0–C_3^0 and the back-transfer produces luminescence. Coupling to other lattice models produces non-radiative pathways in this model. The absence of magnetic or electric field effects originally led to the IVAP model, but stronger support is found in recent optically-detected magnetic resonance data on a-As_2Se_3 [446, 447]. Time resolved luminescence in a-$Si_{1-x}O_x$ alloys has been proposed as evidence of the single centre model [12, 478].

Negative U centres maintain ε_f in the centre of the gap in the presence of extrinsic doping but the conductivity activation energy changes from $E_g/2$ to $E_g/3$ [449–451]. Electrons can be introduced by doping with alkalis which enter interstitial Na^+ sites. Holes are best introduced by doping with group III elements (e.g. Ga), which tend to enter as covalent Ga_4^+ sites. Explaining doping responses is one of the least successful aspects of the VAP model. While changes in mobilities are found, experimentally the photo-induced ESR and luminescence signals of the host remain unchanged in contradiction to the model, this may be related to the formation of un-ionized defect-impurity complexes [452–454].

Two other queries have been recently raised against specific applications of the VAP model. An EHT calculation of As_2Se_3 found substantial σ–π mixing, giving rise to doubts about the ability of its Se sites to overcoordinate [292, 455]. In fact the electronic properties of a-As_2Se_3 provide the best experimental example of a VAP system. The doubts are invalid because EHT calculations overestimate interlayer interactions which caused the mixing, as noted earlier (§ 2.2), and because π states need only exist for overcoordination to be possible.

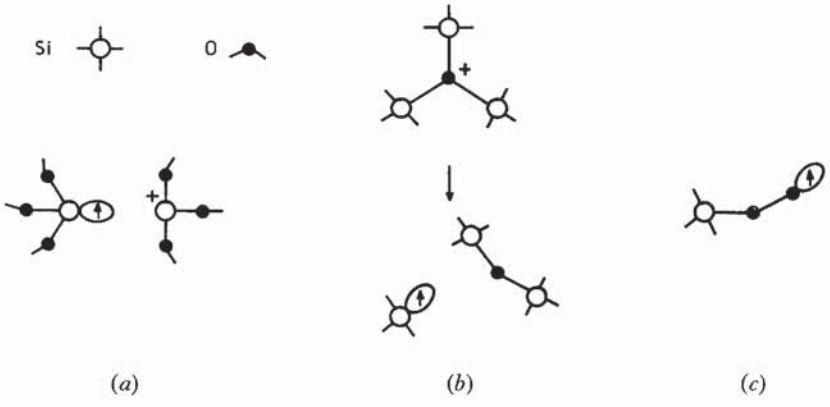
It has also been suggested that a-Se may possess positive U defects based on a calculation of the integrated Hellmann–Feynman force for the $2C_2^0 \rightarrow C_3^+ + C_1^-$ reaction [456]. Paradoxically, a-Se is not the best example of a VAP system and does show certain anomalies, some related to the difficulties with oxygen contamination [442]. However, the experimental evidence in favour of strong electron–lattice coupled defect states is presently too strong to revoke the validity of VAPs in a-Se [434].

8.4. *Defects in a-SiO_2*

SiO_2 is of great technological importance because of its use in optical fibres and as a maskant and passivant in the silicon planar process. The defects of SiO_2 are unique because it is a member of two materials classes, the chalcogenides and the wide-gap insulators. Like other chalcogenides, the twofold co-ordination of oxygen in SiO_2 gives it an O $p\pi$ upper valence band (fig. 31) so VAP defects are anticipated [16, 457–460]. As a wide-gap insulator, radiation can produce electron–hole pairs which can give atomic displacements rather than simple recombination. Thus, vacancies and interstitial-like species are also possible defects in a-SiO_2 [461–463].

In the VAP scheme, the diamagnetic defects are an oxygen surrounded by three silicons, or O_3^+ (Si_3), and the non-bridging oxygen ion, O_1^-. Their paramagnetic daughters are the Si_3^0 and O_1^0 centres (fig. 40). Under irradiation oxygen vacancies are produced and the oxygen is believed to re-enter the network as either protonated non-bridging sites or as peroxyl bridges (fig. 40), rather than remain as an interstitial. The E' centre seen by ESR in neutron and γ-irradiated SiO_2 is accurately modelled by the charged oxygen vacancy, V_0^+ [464, 465]. In it, the remaining non-bonded

Fig. 40



Atomic configurations of defects in a-SiO_2: (*a*) E' centre; (*b*) O_3^+ (Si_3) and the Si_3^0; and (*c*) peroxyl radical.

electron has localized on one of the two Si dangling bonds, and the other Si has relaxed towards the planar configuration adopted by all 3-electron sites. Thus, the paramagnetic silicon defect could be the isolated Si_3^0 dangling bond or the oxygen vacancy, V_0^+. ESR confirms the vacancy model of the E' centre by the absence of the special oxygen sites needed in the VAP model of the Si_3^0/O_3^+ centre and because its ESR g tensor has orthorhombic symmetry rather than the axial symmetry expected for a simple dangling bond [465–467]. Thus, the absence of isolated Si_3^0 centres in ESR appears to cast doubt on the VAP theory for SiO_2. In fact other E'-like centres have been seen in X-irradiated a-SiO_2, which anneal out at low temperatures [468], and which could be these missing Si_3^0 centres. The O_3^+ centres are expected to be more common than Si_3^0, but direct structural evidence for the existence of such diamagnetic centres is always likely to be more elusive.

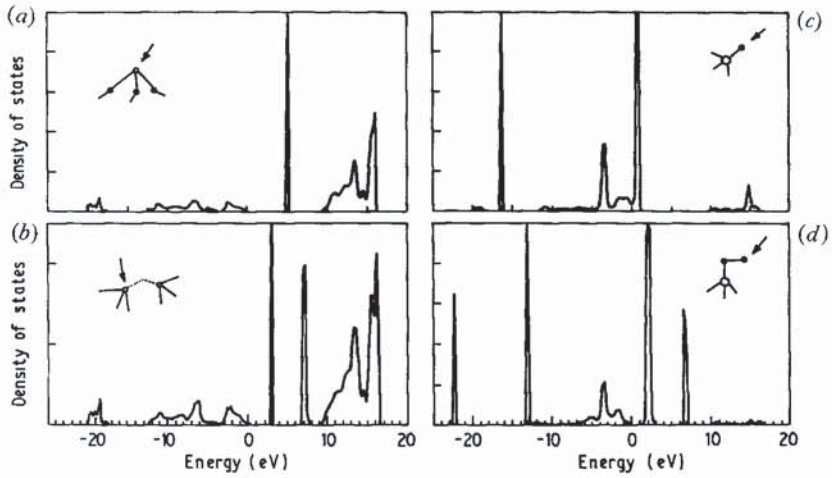
The type of oxygen centre found by ESR depends on the water or hydrogen content of SiO_2 [469–472]. In 'wet' SiO_2 the common oxygen defect is the non-bridging oxygen, O_1^0, associated with a proton. In 'dry' silica the peroxyl radical is the common defect (fig. 40). This is a hole localized primarily on the π orbital of the terminal oxygen. Its suggested parent is the peroxyl bridge, $\equiv Si-O-O-Si\equiv$, which could be described as the Frenkel counterpart of the oxygen vacancy or E' centre.

The absence of Si_3^0 centres in ESR is clearly a problem for a VAP theory of defects in a-SiO_2. The VAP theory postulates a relatively high density of defects in quenched glasses without the need for irradiation [459]. A possible resolution to this conflict is that the VAP creation energy of a-SiO_2 is proportionately higher than in the chalcogenides [340]. A density of VAPs of order $10^{18}\,cm^{-3}$ in normal chalcogenides is attributed to their small creation energy. A higher VAP creation energy of order 2·3 eV could result in only $10^{15}\,cm^{-3}$ defects being quenched into a-SiO_2 at T_g. The greater creation energy could be due to reduced screening in SiO_2 or the planarity of the O_3^+ site, but it does account for the fact that irradiation greatly increases the defect density over the intrinsic level [340].

The levels associated with the various centres were calculated by the TB recursion method (fig. 41) [340]. Si_3^0 dangling bonds were found to give states near the gap centre, 4·9 eV above the valence band. The Si_3^+ centre is higher. The oxygen vacancy has levels related to those of Si_3. For V_0^+ the silicons are well separated and the levels are close to the unperturbed Si_3^+ and Si_3^0 positions. No relaxation of the Si_3^0 sites is expected at V_0^0, so the two Si_3^0 sites are quite close, interact relatively strongly and give strongly localized σ and σ^* states within the gap, characteristic of a weak Si–Si bond [340, 473]. The excitation spectrum of this centre is critically dependent on the strength of the Si–Si interaction and the localization of the σ^* state. The positive charge of the $O_3^+(Si_3)$ sites produces a gap state about 1·5 eV below E_c. The peroxyl bridge and peroxyl radical all give O–O σ^* states in the lower gap region. Additionally, the peroxyl radical and non-bridging oxygen sites give π-like gap states which rise out of the valence band maximum due to self-energy shifts of a similar nature to those found at Se_1^-. These level positions are sufficient to provide a consistent interpretation of many of the optical transitions seen in a-SiO_2 (such as the photoluminescence shown in fig. 39), and confirm the mixed VAP-vacancy character of its defects.

The defects on the a-SiO_2 side of the SiO_2 : Si interface are closely related to those of bulk a-SiO_2. The gaps of SiO_2 and Si line up to give a 3·3 eV conduction band mismatch [474]. Clearly, the excellent passivating ability of SiO_2 means that

Fig. 41



Local DOS at defects in a-SiO_2 (after [340]) with Si—open circle; oxygen—filled circle. (*a*) Si dangling bond; (b) neutral oxygen vacancy; (*c*) oxygen dangling bond; and (*d*) peroxyl radical.

the ideal SiO_2 : Si interface cannot possess interface states. Any traps or ESR centres must be due to defects of some form. ESR has found evidence of Si_3^0 sites on either side of the interface but the SiO_2-side defect near the interface is not the usual E' centre, but possibly an isolated Si_3^0 [475–479]. Such a defect at the interface would differ little from one in bulk SiO_2, but theoretical positioning of its levels with respect to the much narrower gap of Si requires a very accurate treatment.

8.5. *Defects in a-Si*

Much of the current work in amorphous semiconductors is directed to understanding the relationship of defects and doping in a-Si : H. This review considers the problem in four sections (8.5–8.8). Firstly, in evaporated or sputtered a-Si there are of order 10^{20} cm^{-3} gap states and many more reconstructed defect-like states in the band tails [1, 480–482]. Hydrogenation transforms both the 'explicit' and the reconstructed or 'implicit' defect states into passive Si–H states [416, 483, 484]. These states lie outside the gap and allow doping effects to be seen. Hydrogenation leaves a small number of Si dangling bond states in the gap [485–491]. There has been considerable experimental controversy over the density of such states [26, 486, 487, 491, 492]. Finally, recent work has shown that doping of a-Si : H dramatically alters the density of defects, unlike in c-Si, because dopant-defect reactions appear to occur during deposition.

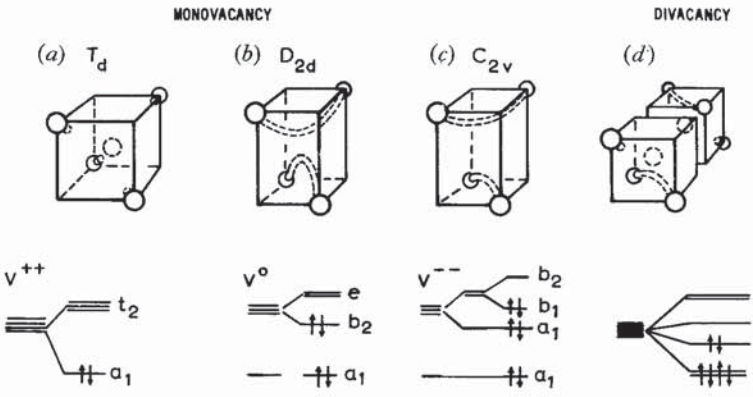
The presence of an ESR signal, variable-range hopping, large a.c. conductivity and a pinned ε_f indicate that a density of gap states of over 10^{20} cm^{-3} exist even in the best a-Si [480–482]. The large distribution of microvoids seen in small angle scattering in a-Si [27] and the 5–10% incorporation of hydrogen in a-Si : H [483] lead one to suspect that the observed gap states are only the tip of the iceberg, and perhaps 5% of all states in a-Si may be defect-related. Unlike a-Se, there is no dominant

defect reconstruction mechanism in a-Si, and we base our understanding on analogies to reconstructions at vacancies and surfaces in c-Si [16].

The ideal monovacancy in c-Si has T_d symmetry (fig. 42). Its four centrally directed dangling bonds form a singly degenerate A_1 state 1 eV below the valence edge and a triply degenerate T_2 gap state. In the lowest occupation V^{2+} only the A_1 state is occupied [437, 493]. Increasing the occupancy places electrons in the degenerate T_2 levels so the vacancy distorts by the Jahn–Teller mechanism. The V^0 state undergoes a D_{2d} tetragonal distortion which partially reforms two bonds across the vacancy. The Jahn–Teller distortion is synonymous with the effective-U mechanism and adjusts the bonding environment to maximize occupation of σ-like states. It is sufficiently strong that the V^{2+}, V^+, V^0 states form a negative U system [437], as does the V^0, V^-, V^{2-} sequence following its C_{2v} distortion. Divacancies are more stable than monovacancies because they only have six dangling bonds. The closest hybrids partially rebond in pairs while the more distant hybrids form a long bond across the divacancy, giving stable non-generate for V^{2+}, V^0 and V^{2-} configurations [494] (the degree of rebonding is unsettled [494 a]).

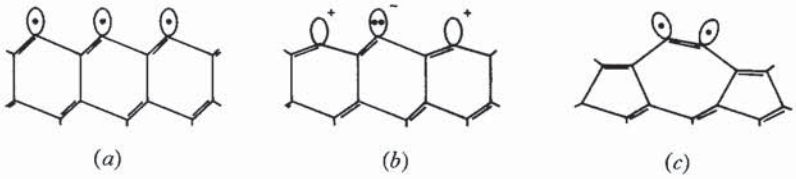
The ideal Si(111) surface has one half-filled dangling bond per site. These are mutual second neighbours, fig. 43 (*a*). Its 2×1 reconstruction is currently debated, between the ionic [495, 496] and π-bonded models [497–500]. Total energy

Fig. 42



Atomic positions and dangling bond occupations at vacancies in c-Si.

Fig. 43



The Si(111) surface: (*a*) Ideal surface; (*b*) ionic reconstruction; and (*c*) π-bonded reconstruction.

calculations, surface band dispersions and core spectra strongly favour the newer, π-bonded model. We discuss the ionic model because of its relevance to the reconstruction of Si(100) surfaces. In the ionic model the surface T_3^0 sites move alternately into and out of the surface, changing their hybridizations and occupations towards sp^2/T_3^+ (cf. B) (fig. 43 (*b*)) and s^2p^3/T_3^- (cf. As), respectively. An ionic gap is introduced in the metallic surface of T_3^0 states. There are two related π-models in which atoms on the two surface layers rearrange topologically to produce rows or islands of nearest-neighbour dangling bonds and, interestingly, five- and sevenfold rings of subsurface bonds [497–499]. The π interaction between the almost parallel dangling bonds opens up a π–π^* gap which is maintained by the low co-ordination within the surface plane.

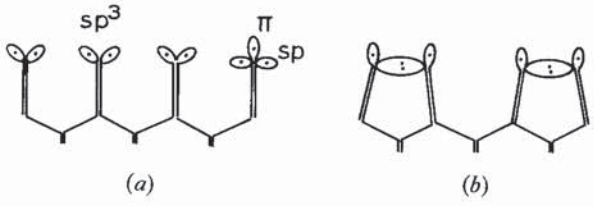
Reconstructions of the Si(100) surface must break the degeneracy of the two dangling bonds per surface site [501–503]. These sp^3 'rabbits ears' can be re-expressed as $p\pi/sp$ hybrids. In the asymmetric dimer model, a 2×1 reconstruction of opposing lateral displacements forms σ bonds between pairs of π orbitals (fig. 44) and alternating vertical diaplacements are needed to split the surface band of sp states, as in the ionic Si(111) 2×1 model. This reconstruction also produces five-fold rings of bonds. We note, however, that the asymmetric dimer model may ultimately need modification in view of the π model of Si(111).

Summarizing these reconstructions in c-Si, we note that the dangling bonds at Si(111) and in the vacancy are all second neighbours, but can be aligned parallel or opposed. Those which are directed towards each other pair up to form weak bonds. Those which are aligned parallel, either form an ionic lattice or drive a more general (π) reconstruction, giving nearest-neighbour dangling bonds which form localized bonding–antibonding levels. Each reconstruction causes a pairing of electrons and may be viewed either as Jahn–Teller distortion or as a negative U effect which breaks the degeneracy at ε_f. There have been a number of direct treatments of defects in a-Si and a-Si : H [13, 504–507]. Total energy calculations suggest that $U > 0$ for isolated T_3^0, that T_3^+/T_3^- sites are only stable as nearest neighbours and that T_2^0 sites are much less stable than combinations of T_3 sites [507].

Two types of defect state can be identified from the preceding discussions, the explicit defect and the implicit defects. Defects are expected at microvoid surfaces in a-Si. Reconstructions expel states from the centre of the gap towards the band tails. This expulsion is driven by a large negative U but once in the tails these implicit states are less localized and have a minimal U. The remaining explicit states deep in the pseudogap produce an ESR signal and have $U > 0$.

The microvoid surfaces are undoubtedly like (111) surfaces of c-Si in that each surface site would usually possess one dangling bond [416]. The c-Si analogy

Fig. 44



The Si(100) surface: (*a*) Ideal surface; (*b*) asymmetric dimer reconstruction.

suggests that these reconstruct according to either the T_3^+/T_3^- model or to give pairs of parallel nearest neighbour T_3^0 sites. The interesting question is the origin of the remaining positive U sites seen in ESR. Connell [131, 364] and others noted that all reconstructions involve pairing and suggested that the deep states result from microvoids starting with odd numbers of dangling bonds. This requires that intervoid communication by bulk bond rearrangements by negligible, but recent calculations suggest that the barrier to such reactions may be small [500]. Phillips suggested that (111)-like surfaces were intrinsic to a-Si and diamagnetic, while the ESR centres correspond to reconstructed (100)-like sites at the surface intersections [416]. The relative number of explicit to implicit defects is then given by the proportion of surface intersections. The Phillips model originated before the π-reconstruction was suggested, but can be generalized to take it into account. We note that (100)-like intersections can convert into adjacent (111)-like defects by bond alteration. We then assume that the surfaces reconstruct but may leave isolated paramagnetic defects because of a clash of phasing during the dimerization of surface sites (reminiscent of the soliton dangling bonds in polyacetylene). This interpretation gives a similar ratio of implicit to explicit defects as in the original Phillips theory.

8.6. *Hydrogen related states in a-Si : H*

Hydrogen passivation of the gap states of a-Si is essential to the observation of doping in a-Si : H [483, 508–514]. The hydrogen content of a-Si : H is at least 6% because both implicit and explicit defects are hydrogenated. Vibrational modes provided much of our early understanding of Si–H bonding and the relative importance of $\equiv Si-H$, $=SiH_2$, $-SiH_3$ groupings [508–511]. Later, electron microscopy and small angle scattering showed how low temperatures and fast deposition rates in the glow discharge (GD) technique produced columnar growth and a second phase with less desirable electronic properties and a high content of $=SiH_2$ groups [508]. NMR revealed that the hydrogen distribution was inhomogeneous even in single phase a-Si : H [512]. a-Si : H can be prepared by sputtering [513]. a-Si : H with low hydrogen content can be prepared by chemical vapour deposition (CVD) at 800°C [515]. A significant ESR signal remains which can be reduced by plasma post-hydrogenation [516]. Finally, low defect density uniform a-Si : H is produced by the low temperature HOMO–CVD process [517, 518]. The gap states of a-Si are also passivated by fluorine in a-Si : H, F [519]. Experimentally, the optical gap of a-Si : H is of order 1·5–2·0 eV, much larger than a-Si or c-Si [520–522]. The increased gap is attributed to a valence band erosion by photoemission, which also finds various H-related states in the valence DOS attributed to different SiH_n configurations [523, 524]. Core absorption [130] show that the conduction band edge of a-Si : H does not move compared to c-Si. Photoemission [4] suggests the gap of a-Si : H extends from $-0{\cdot}5$ to 1·3 eV, compared to the valence edge of c-Si.

The electronic states of Si-H configurations have been studied using pseudopotential, X_α, LCAO and TB calculations on CRNs, model periodic lattices and model clusters [525–532] (fig. 45). The Si–H bond is formed by the interaction of the Si sp^3 hybrid and the H s orbital. Two TB parameterizations of the Si–H bond are available, based on fits to the levels of SiH_4 (table 7) [532, 533].

Experimentally, the principle features of the DOS of a-Si : H containing only monohydride units are the increased gap, the eroded upper valence band, the Si–H σ