Rep. Prog. Phys. **72** (2009) 126501 (29pp) doi:10.1088/0034-4885/72/12/126501

Fundamentals of zinc oxide as a semiconductor

Anderson Janotti and Chris G Van de Walle

Materials Department, University of California, Santa Barbara, CA 93106-5050, USA

E-mail: janotti@engineering.ucsb.edu and vandewalle@mrl.ucsb.edu

Received 10 February 2009, in final form 12 July 2009
Published 22 October 2009
Online at stacks.iop.org/RoPP/72/126501

Abstract

In the past ten years we have witnessed a revival of, and subsequent rapid expansion in, the research on zinc oxide (ZnO) as a semiconductor. Being initially considered as a substrate for GaN and related alloys, the availability of high-quality large bulk single crystals, the strong luminescence demonstrated in optically pumped lasers and the prospects of gaining control over its electrical conductivity have led a large number of groups to turn their research for electronic and photonic devices to ZnO in its own right. The high electron mobility, high thermal conductivity, wide and direct band gap and large exciton binding energy make ZnO suitable for a wide range of devices, including transparent thin-film transistors, photodetectors, light-emitting diodes and laser diodes that operate in the blue and ultraviolet region of the spectrum. In spite of the recent rapid developments, controlling the electrical conductivity of ZnO has remained a major challenge. While a number of research groups have reported achieving p-type ZnO, there are still problems concerning the reproducibility of the results and the stability of the p-type conductivity. Even the cause of the commonly observed unintentional n-type conductivity in as-grown ZnO is still under debate. One approach to address these issues consists of growing high-quality single crystalline bulk and thin films in which the concentrations of impurities and intrinsic defects are controlled. In this review we discuss the status of ZnO as a semiconductor. We first discuss the growth of bulk and epitaxial films, growth conditions and their influence on the incorporation of native defects and impurities. We then present the theory of doping and native defects in ZnO based on density-functional calculations, discussing the stability and electronic structure of native point defects and impurities and their influence on the electrical conductivity and optical properties of ZnO. We pay special attention to the possible causes of the unintentional n-type conductivity, emphasize the role of impurities, critically review the current status of p-type doping and address possible routes to controlling the electrical conductivity in ZnO. Finally, we discuss band-gap engineering using MgZnO and CdZnO alloys.

(Some figures in this article are in colour only in the electronic version)

This article was invited by Professor K Ploog.

Contents

1. Introduction

ZnO is a very promising material for semiconductor device applications [1–5]. It has a direct and wide band gap (figure 1) in the near-UV spectral region [6–10], and a large free-exciton binding energy [6–9] so that excitonic emission processes can persist at or even above room temperature [11, 12]. ZnO crystallizes in the wurtzite structure (figure 1), the same as GaN, but, in contrast, ZnO is available as large bulk single crystals [11]. Its properties have been studied since the early days of semiconductor electronics [13], but the use of ZnO as a semiconductor in electronic devices has been hindered by the lack of control over its electrical conductivity: ZnO crystals are almost always n-type, the cause of which has been a matter of extensive debate and research [1–5]. With the recent success of nitrides in optoelectronics, ZnO has been considered as a substrate to GaN, to which it provides a close match [11]. Over the past decade we have witnessed a significant improvement in the quality of ZnO single-crystal substrates and epitaxial films [1–5]. This, in turn, has led to a revival of the idea of using ZnO as an optoelectronic or electronic material in its own right. The prospect of using ZnO as a complement or alternative to GaN in optoelectronics has driven many research groups worldwide to focus on its semiconductor properties, trying to control the unintentional n-type conductivity and to achieve p-type conductivity. Theoretical studies, in particular first-principles calculations based on density functional theory (DFT), have also contributed to a deeper understanding of the role of native point defects and impurities on the unintentional n-type conductivity in ZnO [14–29]. Acceptor doping has remained challenging, however, and the key factors that would lead to reproducible and stable p-type doping have not yet been identified [1–5].

The availability of large single crystals is a big advantage of ZnO over GaN. For example, GaN is usually grown on sapphire, with a large lattice mismatch of ∼16% that leads to an exceedingly high concentration of extended defects (10^6–10^9 cm^{-2}) [30]. The epitaxy of ZnO films on native substrates can result in ZnO layers with reduced concentration of extended defects and, consequently, better performance in electronic and photonic devices [1–5]. Another big advantage over GaN is that ZnO is amenable to wet chemical etching. This is particularly important in the device design and fabrication.

Band-gap engineering of ZnO can be achieved by alloying with MgO or CdO. Adding Mg to ZnO increases the band gap, whereas Cd decreases the band gap, similar to the effects of Al and In in GaN. Although MgO and CdO crystallize in the rock-salt structure, for moderate concentrations the $Mg_{1-x}Zn_xO$ and $Cd_{1-x}Zn_xO$ alloys assume the wurtzite structure of the parent compound, while still leading to significant band-gap variation.

Controlling the conductivity in ZnO has remained a major issue. Even relatively small concentrations of native point defects and impurities (down to 10^{-14} cm^{-3} or 0.01 ppm) can significantly affect the electrical and optical properties of semiconductors [31–33]. Therefore, understanding the role of native point defects (i.e. vacancies, interstitials, and antisites) and the incorporation of impurities is key toward controlling the conductivity in ZnO. For a long time it has been postulated that the unintentional n-type conductivity in ZnO is caused by the presence of oxygen vacancies or zinc interstitials [34–45]. However, recent state-of-the-art density-functional calculations corroborated by optically detected electron paramagnetic resonance measurements on high-quality ZnO crystals have demonstrated that this attribution to native defects cannot be correct [15, 16, 20, 22, 27, 46, 47]. It has been shown that oxygen vacancies are actually deep donors and cannot contribute to n-type conductivity [20, 46, 47]. In addition, it was found that the other point defects (e.g. Zn interstitials and Zn antisites) are also unlikely causes of the observed n-type conductivity in as-grown ZnO crystals [22, 27].

Instead, the cause would be related to the unintentional incorporation of impurities that act as shallow donors, such as hydrogen which is present in almost all growth and processing environments [14, 26]. By means of density-functional calculations it has been shown that interstitial H forms a strong bond with O in ZnO and acts as a shallow donor, contrary to the amphoteric behavior of interstitial H in conventional semiconductors [14]. Subsequently, interstitial H has been identified and characterized in ZnO [48–50]. However, interstitial H is highly mobile [51, 52] and can easily diffuse out of the samples, making it difficult to explain the stability of the n-type conductivity at relatively high temperatures [53, 54]. More recently, it has been suggested that H can also substitute for O in ZnO and act as a shallow donor [26]. Substitutional H is much more stable than interstitial H and can explain

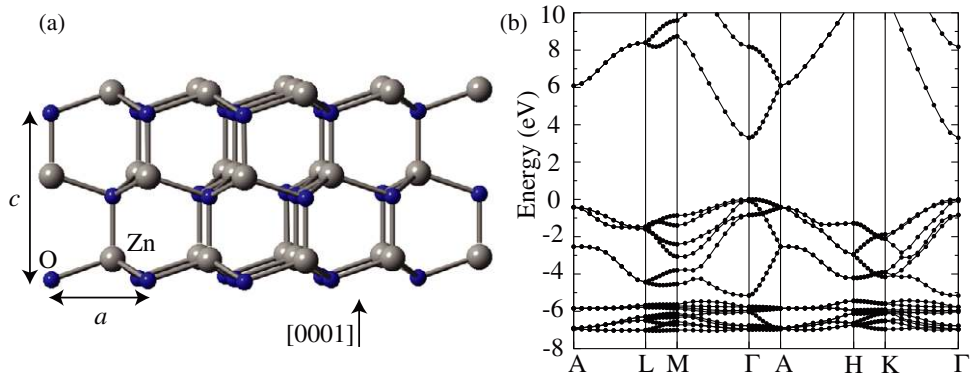

Figure 1. The wurtzite crystal structure of ZnO with the lattice parameters a and c indicated in (a), and the calculated band structure of ZnO using the HSE hybrid functional in (b). The energy of the valence-band maximum (VBM) was set to zero.

the stability of the n-type conductivity and its variation with oxygen partial pressure [26]. Other shallow-donor impurities that emerge as candidates to explain the unintentional n-type conductivity in ZnO are Ga, Al and In. However, these are not necessarily present in all samples in which n-type conductivity has been observed [55].

Obtaining p-type doping in ZnO has proved to be a very difficult task [1–5]. One reason is that ZnO has a tendency toward n-type conductivity, and progress toward understanding its causes is fairly recent [1–5]. Another reason is that the defects, which we now know are *not* responsible for n-type conductivity, *do* play a role as compensating centers in p-type doping [20, 22, 26, 27]. A third reason is the fact that there are very few candidate shallow acceptors in ZnO. Column-IA elements (Li, Na, K) on the Zn site are either deep acceptors or are also stable as interstitial donors that compensate p-type conductivity [56–58]. Column-IB elements (Cu, Ag, Au) are deep acceptors and do not contribute to p-type conductivity. And because O is a highly electronegative first-row element [59], only N is likely to result in a shallow acceptor level in ZnO. The other column-V elements (P, As, Sb) substituting on O sites are all deep acceptors [56]. Quite a few research groups have reported observing p-type conductivity in ZnO [60–69]. In order to explain the reports on p-type doping using P, As or Sb, it was suggested that these impurities would substitute for Zn and form complexes with two Zn vacancies [70]. One problem with this explanation is that these complexes have high formation energies and are unlikely to form. In addition, the reports on p-type ZnO using P, As or Sb often include unexpectedly high hole concentrations, and contain scant information about the crystal quality of the samples or the stability of the p-type conductivity [63–68]. We also note that these reports have not been followed up with reports on stable ZnO p–n junctions. Reports on p-type doping in nitrogen-doped ZnO [62, 69] have provided more detail and display a higher level of consistency. Again, however, they have not been followed up by reports of reproducible p–n junctions, raising questions about the reliability of the observations and the reproducibility and stability of the p-type doping.

A complicating factor in measuring p-type conductivity is the possible formation of a surface electron accumulation layer [71–73]. Under certain conditions, the Fermi level at the ZnO surface may be pinned at surface states located in the conduction band, and an electron accumulation layer may develop near the surface that could severely hinder measurements of the conductivity in the underlying bulk or film. Reports by Schmidt *et al* [71, 72] suggest that the conductivity in ZnO samples is extremely sensitive to the modifications at the surface due to annealing in different environments. Unfortunately, very little is known about surface states in ZnO, and comprehensive investigations on controlled ZnO surfaces still need to be performed in order to assess the possible formation of a surface electron accumulation layer and its effects on electrical measurements. It is also worth noting that Hall-effect measurements in ZnO seem to be particularly prone to misinterpretation, potentially even yielding the wrong carrier type [74, 75]. As recently pointed out by Bierwagen *et al* [75], wrong conclusions about carrier type can result if inhomogeneities are present in the sample. Judicious placement of contacts in van der Pauw/Hall-effect experiments is essential. It has been found that inhomogeneities in carrier mobility do not affect the measured carrier type, as long as the carrier concentration remains homogeneous. However, lateral inhomogeneities in carrier concentrations can result in an incorrect assignment of the carrier type. Problems can be avoided if contacts are placed at the sample corners (for example, in the case of a square sample) and not in the interior of the sample area [75]. Correct placement of the contacts in Hall measurements yields qualitatively correct results even in samples with inhomogeneous mobility and carrier concentration. In this case the measured carrier concentration will be close to the average carrier concentration in the sample [75].

In the following sections we discuss in depth each of the above-raised issues related to ZnO as a semiconductor. In section 2 we describe the physical properties of ZnO and relate them to current or envisioned applications in electronic and optoelectronic devices. In section 3 we give a brief description of the techniques used to grow ZnO, and discuss the quality of ZnO single-crystal substrates and epitaxial films, with emphasis on the electrical properties and background impurity concentrations. In section 4 we discuss in detail the theory of native point defects in ZnO, based on first-principles

density-functional calculations. We describe the electronic structure and local lattice relaxations of all native defects, their formation energies and stability, and emphasize the relation of these results to experimental observations. In particular we discuss the role of point defects on n-type and p-type doping. In sections 5 and 6 we discuss the electronic and structural properties of the most relevant donor and acceptor impurities in ZnO. We describe the role of hydrogen in some detail, and, in particular, the current status of p-type doping in ZnO. In section 7 we briefly review the results for ZnO-based alloys and discuss the deformation potentials and band alignments of MgZnO and CdZnO alloys, based on the properties of the parent compounds ZnO, MgO and CdO. These quantities are important ingredients in the design of optoelectronic devices based on heterointerfaces and quantum wells. Finally, in section 8 we comment on the future of ZnO as a semiconductor.

2. Properties and device applications

The wide range of useful properties displayed by ZnO has been recognized for a long time [13]. What has captured most of the attention in recent years is the fact that ZnO is a semiconductor with a direct band gap of 3.44 eV [7–9], which in principle enables optoelectronic applications in the blue and UV regions of the spectrum. The prospect of such applications has been fueled by impressive progress in bulk-crystal [76–78] as well as thin-film growth over the past few years [62, 79–83]. A partial list of the properties of ZnO that distinguish it from other semiconductors or oxides or render it useful for applications includes:

- *Direct and wide band gap.* The band gap of ZnO is 3.44 eV at low temperatures and 3.37 eV at room temperature [7]. For comparison, the respective values for wurtzite GaN are 3.50 eV and 3.44 eV [84]. As mentioned above, this enables applications in optoelectronics in the blue/UV region, including light-emitting diodes, laser diodes and photodetectors [1–5]. Optically pumped lasing has been reported in ZnO platelets [11], thin films [12], clusters consisting of ZnO nanocrystals [85] and ZnO nanowires [86]. Reports on p–n homojunctions have recently appeared in the literature [69, 87–89], but stability and reproducibility have not been established.
- *Large exciton binding energy.* The free-exciton binding energy in ZnO is 60 meV [11, 12], compared with, e.g. 25 meV in GaN [84]. This large exciton binding energy indicates that efficient excitonic emission in ZnO can persist at room temperature and higher [11, 12]. Since the oscillator strength of excitons is typically much larger than that of direct electron–hole transitions in direct gap semiconductors [90], the large exciton binding energy makes ZnO a promising material for optical devices that are based on excitonic effects.
- *Large piezoelectric constants.* In piezoelectric materials, an applied voltage generates a deformation in the crystal and vice versa. These materials are generally used as sensors, transducers and actuators. The low symmetry of the wurtzite crystal structure combined with a large electromechanical coupling in ZnO gives rise to strong piezoelectric and pyroelectric properties. Piezolectric ZnO films with uniform thickness and orientation have been grown on a variety of substrates using different deposition techniques, including sol–gel process, spray pyrolysis, chemical vapor deposition, molecular-beam epitaxy and sputtering [91–98].
- *Strong luminescence.* Due to a strong luminescence in the green–white region of the spectrum, ZnO is also a suitable material for phosphor applications. The emission spectrum has a peak at 495 nm and a very broad half-width of 0.4 eV [99]. The n-type conductivity of ZnO makes it appropriate for applications in vacuum fluorescent displays and field emission displays. The origin of the luminescence center and the luminescence mechanism are not really understood, being frequently attributed to oxygen vacancies or zinc interstitials, without any clear evidence [99]. As we will discuss later, these defects cannot emit in the green region, and it has been suggested that zinc vacancies are a more likely cause of the green luminescence. Zn vacancies are acceptors and likely to form in n-type ZnO.
- *Strong sensitivity of surface conductivity to the presence of adsorbed species.* The conductivity of ZnO thin films is very sensitive to the exposure of the surface to various gases. It can be used as a cheap smell sensor capable of detecting the freshness of foods and drinks, due to the high sensitivity to trimethylamine present in the odor [100]. The mechanisms of the sensor action are poorly understood. Recent experiments reveal the existence of a surface electron accumulation layer in vacuum annealed single crystals, which disappears upon exposure to ambient air [71–73]. This layer may play a role in sensor action, as well. The presence of this conducting surface channel has been suggested to be related to some puzzling type-conversion effects observed when attempting to obtain p-type ZnO [71–73].
- *Strong non-linear resistance of polycrystalline ZnO.* Commercially available ZnO varistors are made of semiconducting polycrystalline films with highly non-ohmic current–voltage characteristics. While this non-linear resistance has often been attributed to grain boundaries, the microscopic mechanisms are still not fully understood and the effects of additives and microstructures, as well as their relation to degradation mechanisms, are still under debate [101].
- *Large non-linear optical coefficients.* ZnO crystals and, in particular, thin films exhibit second- and third-order non-linear optical behavior, suitable for non-linear optical devices. The linear and non-linear optical properties of ZnO depend on the crystallinity of the samples. ZnO films grown by laser deposition, reactive sputtering and spray pyrolysis show strong second-order non-linear response. Third-order non-linear response has recently been observed in ZnO nanocrystalline films [102]. The non-linear optical response in ZnO thin films is attractive for integrated non-linear optical devices.
- *High thermal conductivity.* This property makes ZnO useful as an additive (e.g. ZnO is added to rubber in

order to increase the thermal conductivity of tires). It also increases the appeal of ZnO as a substrate for homoepitaxy or heteroepitaxy (e.g. for growth of GaN, which has a very similar lattice constant) [103, 104]. High thermal conductivity translates into high efficiency of heat removal during device operation.

- *Availability of large single crystals.* One of the most attractive features of ZnO as a semiconductor is that large area single crystals are available, and epi-ready substrates are now commercialized. Bulk crystals can be grown with a variety of techniques, including hydrothermal growth [77, 105, 106], vapor-phase transport [76] and pressurized melt growth [107, 108]. Growth of thin films can be accomplished using chemical vapor deposition (MOCVD) [82, 83], molecular-beam epitaxy [80, 81], laser ablation [109] or sputtering [110]. The epitaxial growth of ZnO on native substrates can potentially lead to high-quality thin films with reduced concentrations of extended defects. This is especially significant when compared with GaN, for which native substrates do not exist. In view of the fact that the GaN-based devices have achieved high efficiencies despite the relatively large concentration of extended defects, it is possible that a high-quality ZnO-based device could surpass the efficiencies obtained with GaN.
- *Amenability to wet chemical etching.* Semiconductor device fabrication processes greatly benefit from the amenability to low-temperature wet chemical etching. It has been reported that ZnO thin films can be etched with acidic, alkaline as well as mixture solutions. This possibility of low-temperature chemical etching adds great flexibility in the processing, designing and integration of electronic and optoelectronic devices.
- *Radiation hardness.* Radiation hardness is important for applications at high altitude or in space. It has been observed that ZnO exhibits exceptionally high radiation hardness [111, 112], even greater than that of GaN, the cause of which is still unknown.

In addition to the above-mentioned properties and applications it is worth mentioning that, similarly to GaN-based alloys (InGaN and AlGaN), it is possible to engineer the band gap of ZnO by adding Mg and/or Cd. Although CdO and MgO crystallize in the rock-salt structure, for moderate concentrations MgZnO and CdZnO assume the wurtzite structure of ZnO with band gaps in the range of 2.3 to 4.0 eV [113–118]. It is also worth noting that ZnO substrates offer a perfect lattice match to $In_{0.22}Ga_{0.78}N$, which has a band gap highly suitable for visible light emission. ZnO has also attracted attention due to the possibility of making thin-film transistors on flexible substrates with relatively high electron mobility when compared with amorphous silicon or organic semiconductors [119–121].

In the following we discuss the current status of the growth of ZnO substrates and thin films. We focus on the quality aspects that are related to the levels of background n-type conductivity and impurity incorporation.

3. Growth of ZnO bulk and epitaxial films

For most of its current applications ZnO is used in the polycrystalline form, and crystalline quality or purity is not an issue. For more advanced applications, single crystals in the form of bulk or thin films and a high degree of purity are required. Several groups have pursued growth of ZnO thin films and bulk, and the rapid progress in improving quality and purity is impressive. Bulk crystals with size up to 2 inches have been obtained and films grown on ZnO (homoepitaxy) or other substrates (heteroepitaxy) have been obtained. Despite the rapid progress, a more detailed understanding of homoepitaxy is necessary. Homoepitaxy was, at first, thought to be straightforward, but has been found to be far from straightforward. In the following we discuss bulk and epitaxial film growth, the common impurities found in these materials and the crystalline quality, electrical and optical properties.

3.1. Bulk growth

Growth of zinc oxide bulk can be carried out by a variety of methods, including gas or vapor transport, hydrothermal and pressurized melt growth. These techniques involve different growth mechanisms, resulting in bulk crystals grown at different rates, with different impurity background concentrations and, consequently, different electrical and optical properties.

In the gas-transport technique, one usually starts with purified ZnO powder that is reduced to Zn vapor at elevated temperatures (∼1600 K) by hydrogen or graphite. The zinc vapor is then oxidized in a region of low temperature under oxygen or air, resulting in ZnO platelets or hexagonal needles with diameters up to several millimeters and lengths of several centimeters [122–125], as shown in figure 2(*a*). In the seeded vapor transport method, ZnO powder is used as the ZnO source at the hot end of a horizontal tube held at temperatures above 1150 °C. Transport of material to the cooler end of the tube proceeds by using a carrier gas (e.g. H_2). Assisted by a single-crystal seed, bulk ZnO is then formed at the cool end of the tube. The state-of-the-art seeded chemical vapor transport (SCVT) technique produces ZnO single crystals 2 inches in diameter and 1 cm in thickness in about 150 h with a growth rate of 1 mm day^{-1} [76]. The SCVT ZnO samples are also n-type, with a typical room temperature carrier concentration of $\sim 10^{16}$ cm^{-3}. Room temperature mobility of 205 cm^2 V^{-1} s^{-1} and a peak mobility of ∼2000 cm^2 V^{-1}s^{-1} at 50 K have been reported [76]. The estimated concentration of the dominant donor is about 10^{17} cm^{-3} and the total concentration of acceptors is about 10^{15} cm^{-3}. Peaks in the low-temperature photoluminescence (PL) spectrum indicate the presence of more than one type of donor, and the broad green band is a factor of 4000 weaker than the band-edge emission.

In the hydrothermal method, the growth takes place in a platinum-lined autoclave held at relatively low temperatures in the range 300–400 °C. ZnO is dissolved in a KOH/LiOH base solution in a high temperature and pressure region, and

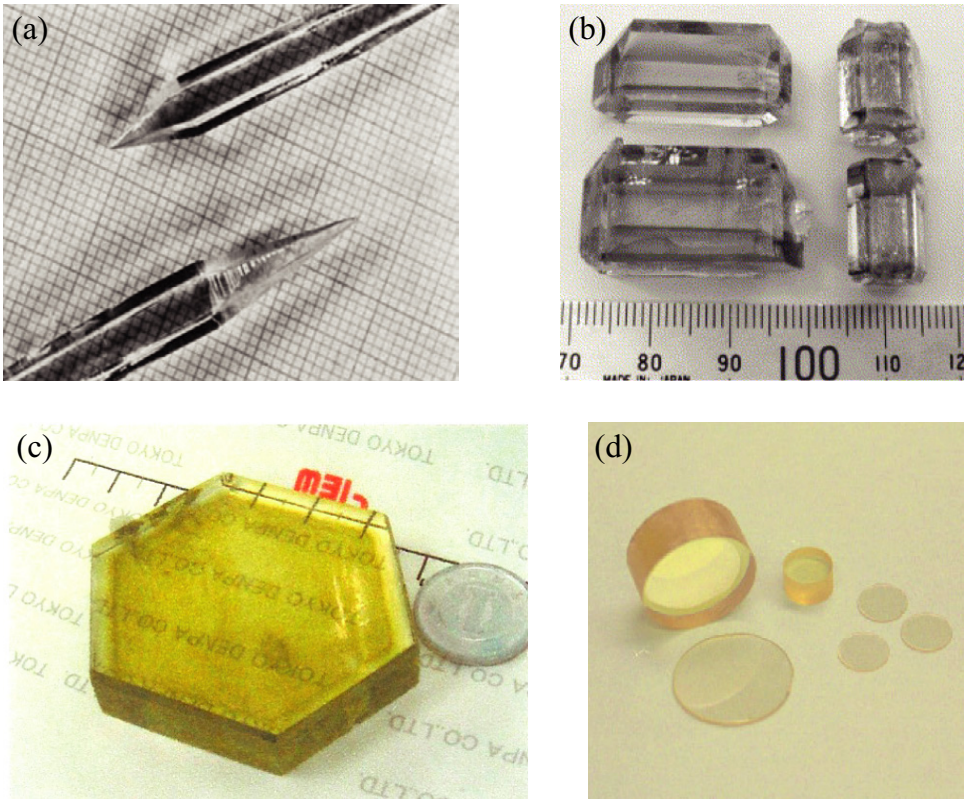

Figure 2. Photographs of large bulk ZnO single crystals grown by different techniques: (*a*) gas transport, (*b*) hydrothermal, (*c*) hydrothermal and (*d*) pressurized melt growth. From [77, 106, 108, 125].

precipitated in a region of reduced temperature. Hydrothermal growth has resulted in large ZnO crystals with high crystalline quality up to 2 inches in diameter (figure 2(*b*) and (*c*)), allowing the production of high-quality large substrates for homoepitaxy and heteroepitaxy [77, 106, 126]. Hydrothermal ZnO requires relatively low growth temperatures compared with the other methods. It is characterized by slow growth rates of about 0.03 inches day^{-1} and unavoidable incorporation of impurities coming from the solvent, such as Li and K, that may strongly affect the electrical properties of these type of samples. Maeda *et al* [77] reported the presence of Li and K in concentrations of 0.9 ppm and 0.3 ppm, respectively, accompanied by lower concentrations of Al and Fe. The incorporation of Li ($\sim 10^{16}$ cm^{-3}) is probably related to the low electron concentration (8×10^{13} cm^{-3}) and high resistivity (380 Ω cm) in hydrothermal ZnO [77]. Li on a Zn site is a deep acceptor that compensates the n-type conductivity caused by other impurities. The PL spectrum at 11 K shows a strong and sharp emission around 3.4 eV and a much weaker and broad band in the green region (~2.4 eV). It has also been observed that annealing the ZnO samples at 1100 °C for 4 h under 1 atm significantly reduces the etch pit density from 300 to 80 cm^{-2}, dramatically improving the surface morphology [77].

Large ZnO bulk crystals have also been grown from the melt, through a pressurized melt-growth technique patented by Cermet, Inc. [108]. In this modified Bridgman process, radio-frequency energy is used as a heat source to produce a molten phase in a cold-wall crucible, in a controlled gas atmosphere. The ZnO single crystal is isolated from the crucible by a cooled ZnO layer, thus reducing impurity contamination from the crucible. The technique allows for obtaining ZnO boules which are 1 cm in diameter and several centimeters thick in much less time (1–5 mm h^{-1}) than the hydrothermal and seeded vapor transport methods. Melt-grown ZnO crystals can then be cut into epitaxial-ready oriented wafers [107, 108]. Melt-grown ZnO is also of high crystalline quality, with a reduced concentration of extended defects on the order of 10^4 cm^{-2}. The low-temperature PL spectrum reveals a large number of exciton lines near a sharp band-edge emission. A typically weaker and broad green band emission is also observed. The Cermet samples show high unintentional n-type conductivity, with carrier concentrations on the order of 10^{17} cm^{-3} and carrier mobility of ~130 cm^2 V^{-1}s^{-1} at room temperature [107, 108].

Note that the as-grown ZnO bulk single crystals are always n-type irrespective of the growth method. The cause of this unintentional n-type conductivity has been widely discussed in the literature, and has often been attributed to the presence of native point defects such as oxygen vacancies and zinc interstitials. However, recent first-principles calculations indicate that oxygen vacancy is actually a deep donor, and cannot contribute to the observed n-type conductivity. The calculated optical transitions related to oxygen vacancies agree very well with optically detected electron paramagnetic resonance (ODEPR) measurements, confirming the deep donor character of oxygen vacancy in ZnO. Moreover, the ODEPR signals related to oxygen vacancies are not observed in the as-received (as-grown) ZnO bulk samples grown by SCVT

or from the melt, but only after irradiation with high-energy electrons. This indicates that oxygen vacancies are not present in as-grown bulk ZnO. In addition, first-principles calculations indicate that zinc interstitials are shallow donors, but unstable; they have high formation energies and very low migration barriers, in agreement with experimental results of irradiated samples. As we will discuss in sections 4 and 5, it is likely that the n-type conductivity observed in as-grown bulk ZnO single crystals is caused by the unintentional incorporation of impurities, with H being a plausible candidate since it is difficult to avoid its presence in most growth and annealing environments.

Another common feature observed in as-grown ZnO bulk single crystals is the presence of a weak and broad green band in the PL spectrum. The cause of the green emission in ZnO has also been widely debated in the literature. Recently, it has been suggested that the zinc vacancy is a major cause. As we will discuss later, zinc vacancies are indeed deep acceptors and likely to be present in n-type samples. Experiments have also indicated that zinc vacancies are the dominant native point defects present in as-grown Zn bulk crystals [127].

3.2. Epitaxial thin-film growth

The main advantage of having high-quality large single crystals of ZnO available is that ZnO thin films or layers can in principle be epitaxially grown with reduced concentrations of extended defects, without contamination from the substrate, and without a thermal mismatch. This is especially important for optoelectronic devices in which the performance is highly sensitive to the crystalline quality of the layers.

Although ZnO substrates have been available for a long time, most ZnO epitaxial layers have been grown on non-native substrates including sapphire, GaAs, CaF_2, $ScAlMgO_4$, Si and GaN [65, 82, 87, 128–137], with only a few reports on homoepitaxial growth of ZnO layers [80, 83, 138, 139]. This can be attributed, to some extent, to the current high price of ZnO substrates, and also to insufficient knowledge about appropriate surface preparation for epitaxy. It has recently been reported that ZnO surfaces have to be carefully treated prior to epitaxy in order to avoid the tendency toward columnar or 3D growth that results in rough surface morphology [83, 137].

Most of the current technological applications of ZnO, such as varistors, transparent conductive electrodes for solar cells, piezoelectric devices and gas sensors, have made use of polycrystalline films that are grown by a variety of deposition techniques, mostly on glass substrates. These techniques include chemical spray pyrolysis, screen painting, electrochemical deposition, sol–gel synthesis and oxidation of Zn films, and are characterized by requiring relatively low temperatures and covering large areas. However, we emphasize that for electronic and optoelectronic applications, high-quality single-crystal epitaxial films with minimal concentrations of native defects and controlled impurity incorporation are required. For these, optimized growth and processing environments (partial pressures and temperature) are necessary. Current techniques that allow for this level of control include pulsed laser deposition (PLD), chemical vapor deposition (CVD), metal-organic CVD (MOCVD) and molecular-beam epitaxy (MBE), and to a lesser extent sputtering. Magnetron sputtering is recognized to be the most scalable technique, at the expense of lower crystalline quality, often resulting in columnar structures. The lower crystalline quality of the ZnO films grown by sputtering techniques likely arises from the difficulties in controlling particles landing on the film surface, preventing the growth of defect-free films with good optical quality [143, 144]. In the following we briefly describe the results for ZnO thin films grown by PLD, MOCVD and MBE techniques, focusing on the crystalline quality, electrical and optical properties and background impurity incorporation. More extensive discussions on epitaxial growth of ZnO are available in the literature [2, 3, 5, 79, 140–143].

In the PLD method a high-power laser beam is focused inside a chamber to strike a target of known composition, producing a highly directed plume of gas material which condenses onto a substrate [142, 143]. Targets used for growing ZnO films by PLD are sintered ceramic disks prepared from high-purity pressed powders, ZnO single crystals or pure Zn with a reactive oxygen atmosphere. MgZnO and CdZnO alloys and doping can be achieved by either including the alloying elements and dopants in the target or using a reactive gas in the chamber. Glass substrates as well as single-crystal substrates have been used to grow ZnO thin films using PLD, with the best results obtained using the latter. Sapphire has been the most used substrate due to the large area of the single-crystal wafers and the low cost. Other single-crystal substrates have also been used to grow ZnO by PLD, including Si, GaAs, InP, CaF_2 and $LiTaO_3$. However, most of these substrates have a large lattice mismatch with ZnO, and the deposited films contain large-size crystallites separated by grain boundaries that are detrimental to semiconductor applications. The relatively low Hall mobility of less than $160\,cm^{-3}$ observed in ZnO films grown by PLD is attributed to the dominance of carrier scattering at grain boundaries [142, 143]. Recent results on PLD ZnO films grown on $ScAlMgO_4$ (SCAM) deserve special attention. SCAM has a relatively small lattice mismatch of 0.09% with ZnO, and has proved the best alternative to sapphire substrates. ZnO films grown on $ScAlMgO_4$ have shown high crystal quality, low defect densities and high Hall mobility of $440\,cm^2\,V^{-1}\,s^{-1}$ [129].

MOCVD and MBE are expected to lead to better ZnO films in terms of crystalline quality, yet at the expense of slow growth rates and much more complicated setups. In MOCVD, the epitaxial layer grows via chemical reactions of the constituent chemical species at or near the heated substrate [80, 82, 83, 138]. In contrast, in MBE the epitaxial films grow by physical deposition. MOCVD takes place in gas phase at moderate pressures, and has become the preferred technique for the growth of devices and the dominant process for the manufacture of laser diodes, solar cells and LEDs. Very promising results have already been obtained in the MOCVD growth of ZnO films, with the best layers obtained by homoepitaxy as expected. MOCVD ZnO films have been grown on a wide range of substrates including glass, sapphire, Si, Ge, GaAs, GaP, InP, GaN and ZnO,

with electron concentrations varying from 10^{15} to $10^{20}\,cm^{-3}$. Only a few studies have been performed on MBE growth of ZnO epitaxial layers, with the first report released in 1996 [145]. Substrates include sapphire, $LiTaO_3$, MgO and GaN. Electron concentrations from 10^{16} to $10^{18}\,cm^{-3}$, and mobilities in the range 90–260 $cm^2\,V^{-1}s^{-1}$ have been reported.

3.3. Conductivity control

Note that most of the growth techniques produce ZnO that is highly n-type. This high level of n-type conductivity is very useful for some applications, such as transparent conductors—but in general it would be desirable to have better control over the conductivity. In particular, the ability to reduce the n-type background and to achieve p-type doping would open up tremendous possibilities for semiconductor device applications in general and for light-emitting diodes and lasers in particular.

Hence, controlling the n-type conductivity in ZnO is a topic of much interest. Much of the debate still surrounding this issue is related to the difficulties in unambiguously detecting the actual cause of doping, and distinguishing between point defects and impurities as the source. We will focus on some of these issues later in the text; for now, we point out the following complications:

(i) Unintentional incorporation of impurities is very difficult or even impossible to exclude. Impurities are introduced from sources or precursors (gaseous or solid), they can diffuse out of the substrate, or they can emanate from the walls of the growth chamber. Even in the ultrahigh vacuum environment used in MBE, the background concentration of residual gases (mostly hydrogen) is high enough so that incorporation of a high-solubility contaminant cannot be excluded. Until the 1990s, quantitative measurement techniques to assess impurity concentrations down to the ppm range were either not available or not widely used. The use of secondary-ion mass spectrometry (SIMS) has had a huge impact.

(ii) Measurements of stoichiometry are even more difficult than measurements of impurity concentrations. While the latter can be compared with looking for the proverbial needle in a haystack, assessing stoichiometry requires identifying the presence (or absence) of an extra sprig of hay itself. Even if accurate data are available, it is by no means certain that the deviation from stoichiometry is accommodated through the formation of point defects, as opposed to clusters, precipitates or extended defects (such as grain boundaries or dislocations).

(iii) Attributions to point defects have often been made on the basis of observed changes in conductivity as a function of oxygen partial pressure. But changes in partial pressure can have a number of simultaneous effects. For instance, a decrease in oxygen pressure could make it more likely that oxygen vacancies (V_O) are formed in ZnO; however, when hydrogen is present, it also becomes more likely that hydrogen can incorporate on oxygen sites (H_O). Since H_O acts as a shallow donor in ZnO (see section 5), a correlation between a change in conductivity and a change in oxygen partial pressure does not unambiguously identify oxygen vacancies as the source of conductivity. It can explain, however, the historic tendency of attributing the often-observed n-type conductivity in ZnO to the presence of oxygen vacancies.

Besides controlling the n-type conductivity in ZnO epitaxial layers, the biggest challenge in research on ZnO as a semiconductor is to achieve p-type doping. There are in fact numerous reports on p-type doping in the literature, with hole concentrations varying from $10^{16}\,cm^{-3}$ to values as high as $10^{18}\,cm^{-3}$, and hole mobility varying from 0.1 to 50 $cm^2\,V^{-1}\,s^{-1}$ [62, 65, 130–136]. However, reliability and reproducibility are still big issues, and the interpretation of the results has been controversial [146]. No reliable devices based on p–n homojunction have been reported so far.

4. Native point defects in ZnO

Native or intrinsic defects are imperfections in the crystal lattice that involve only the constituent elements [31]. They include vacancies (missing atoms at regular lattice positions), interstitials (extra atoms occupying interstices in the lattice) and antisites (a Zn atom occupying an O lattice site or vice versa). Native defects can strongly influence the electrical and optical properties of a semiconductor, affecting doping, minority carrier lifetime and luminescence efficiency, and are directly involved in the diffusion mechanisms connected to growth, processing and device degradation [31–33]. Understanding the incorporation and behavior of point defects in ZnO is therefore essential to its successful application in semiconductor devices.

Native defects are, in general, related to the compensation of the predominant acceptor or donor dopants, i.e. donor defects are easier to form in p-type material, whereas acceptor defects are easier to form in n-type material, always counteracting the prevailing conductivity. Native defects have long been believed to play an even more important role in ZnO, which frequently exhibits high levels of unintentional n-type conductivity. Oxygen vacancies and zinc interstitials have often been invoked as sources of n-type conductivity in ZnO [34–45]. However, most of these arguments are based on indirect evidence, e.g. that the electrical conductivity increases as the oxygen partial pressure decreases. In our view, these statements about the role of native point defects as sources of conductivity are only hypotheses that are not supported by experimental observations. In fact, they are in contradiction with several careful experiments, as well as with accurate density-functional calculations. In the following we discuss the theory of point defects in ZnO, with an emphasis on results of density-functional calculations, and relate it to the experimental observations whenever possible.

4.1. Defect concentrations and formation energies

Assuming thermodynamic equilibrium and neglecting defect–defect interactions (i.e. in the dilute regime), the concentration

of a native defect in a solid is determined by its formation energy E^{f} through the relation [147]

$$c = N_{\mathrm{sites}} \exp\left(\frac{-E^{\mathrm{f}}}{k_{\mathrm{B}}T}\right), \quad (1)$$

where N_{sites} is the number of sites (including different configurations) per unit volume the defect can be incorporated on, k_{B} is the Boltzmann constant and T the temperature. Equation (1) shows that defects with high formation energies will occur in low concentrations. The energy appearing in equation (1) is, in principle, a *free energy* of formation; however, contributions from the formation volume and the formation entropy are often neglected since they are small or negligible at the relevant experimental conditions. The formation volume is related to the change in the volume when the defect is introduced into the system, being negligible in the dilute regime; it tends to become important only at very high pressures. The formation entropy is related mainly to the change in the frequency of the vibrational modes of the crystal containing the defect with respect to the perfect crystal—note that the configurational entropy is already included in the derivation of equation (1) [147]. Formation entropies of point defects are typically of the order of a few k_{B}, and therefore their contribution to the free energy is much smaller than formation energies which are on the order of 1 eV or more [148]. In addition, significant cancellation effects usually occur. Therefore, even at high temperatures the contributions from formation entropies are usually small.

The formation energy of a defect or impurity and, hence, its concentration can be computed entirely from first principles, without resorting to experimental data. Density-functional theory allows us to calculate the ground-state total energy of systems of electrons subject to an external potential, i.e. the Coulomb potential given by the nuclei or ions. From total energies one can easily compute the formation energy of defects [148, 149] as described below. First-principles density-functional calculations of defects in solids are nowadays performed using supercells containing up to several 100 atoms, periodically repeated in the three-dimensional space. Obviously the supercell size is limited by the computational cost, but it should be large enough to simulate isolated defects, i.e. the interactions between defects in neighboring supercells should be small. In the following we describe the calculation of formation energies, transition levels and migration barriers.

The formation energy of a point defect depends on the growth or annealing conditions [148]. For example, the formation energy of an oxygen vacancy is determined by the relative abundance of Zn and O atoms in the environment, as expressed by the chemical potentials μ_{Zn} and μ_{O}, respectively. If the vacancy is charged, the formation energy further depends on the Fermi level (E_{F}), which is the energy of the electron reservoir, i.e. the electron chemical potential. In the case of an oxygen vacancy in ZnO, the formation energy is given by

$$E^{\mathrm{f}}(V_{\mathrm{O}}^{q}) = E_{\mathrm{tot}}(V_{\mathrm{O}}^{q}) - E_{\mathrm{tot}}(\mathrm{ZnO}) + \mu_{\mathrm{O}} + q(E_{\mathrm{F}} + E_{\mathrm{VBM}}), \quad (2)$$

where $E_{\mathrm{tot}}(V_{\mathrm{O}}^{q})$ is the total energy of a supercell containing the oxygen vacancy in the charge state q, $E_{\mathrm{tot}}(\mathrm{ZnO})$ is the total energy of a ZnO perfect crystal in the same supercell and μ_{O} is the oxygen chemical potential. Expressions similar to equation (2) apply to all native point defects.

The chemical potential μ_{O} can be related to the experimental conditions, which can be either Zn-rich, O-rich or anything in between, and is, therefore, explicitly regarded as a variable in the formalism. However, the thermodynamic equilibrium and the stability of ZnO impose bounds on the chemical potential. The oxygen chemical potential μ_{O} is subject to an upper bound given by the energy of O in an O_2 molecule, $\mu_{\mathrm{O}}^{\mathrm{max}} = 1/2E_{\mathrm{tot}}(O_2)$, corresponding to extreme O-rich conditions. Similarly, the zinc chemical potential μ_{Zn} is subject to an upper bound given by the energy of Zn in bulk zinc, $\mu_{\mathrm{Zn}}^{\mathrm{max}} = E_{\mathrm{tot}}(\mathrm{Zn})$, corresponding to extreme Zn-rich conditions. It should be kept in mind that μ_{O} and μ_{Zn}, which are free energies, are temperature and pressure dependent.

The upper bounds defined above also lead to lower bounds given by the thermodynamic stability condition for ZnO, i.e.

$$\mu_{\mathrm{Zn}} + \mu_{\mathrm{O}} = \Delta H_{\mathrm{f}}(\mathrm{ZnO}), \quad (3)$$

where $\Delta H_{\mathrm{f}}(\mathrm{ZnO})$ is the enthalpy of formation of bulk ZnO (which is negative for a stable compound). The upper limit on the zinc chemical potential then results in a lower limit on the oxygen chemical potential, $\mu_{\mathrm{O}}^{\mathrm{min}} = 1/2E_{\mathrm{tot}}(O_2) + \Delta H_{\mathrm{f}}(\mathrm{ZnO})$. Conversely, the upper limit on the oxygen chemical potential results in a lower limit on the zinc chemical potential, $\mu_{\mathrm{Zn}}^{\mathrm{min}} = E_{\mathrm{tot}}(\mathrm{Zn}) + \Delta H_{\mathrm{f}}(\mathrm{ZnO})$. Enthalpies of formation calculated from first principles are usually quite accurate. For ZnO, $\Delta H_{\mathrm{f}}(\mathrm{ZnO}) = -3.5$ eV, compared with the experimental value of -3.6 eV [150]. This value indicates that the chemical potentials μ_{O} and μ_{Zn} and, consequently, the defect formation energies can in principle vary over a wide range, corresponding to the magnitude of the enthalpy of formation of ZnO.

The Fermi level E_{F} in equation (2) is not an independent parameter, but is determined by the condition of charge neutrality. In principle equations such as (2) can be formulated for every native defect and impurity in the material; the complete problem, including free-carrier concentrations in valence and conduction bands, can then be solved self-consistently by imposing the charge neutrality condition. However, it is instructive to plot formation energies as a function of E_{F} in order to examine the behavior of defects when the doping level changes. The Fermi level E_{F} is taken with respect to the valence-band maximum (VBM), and can vary from 0 to E_{g}, where E_{g} is the fundamental band gap. Note that E_{VBM} in equation (2) is taken from a calculation for the perfect crystal, corrected by aligning the averaged electrostatic potential in the perfect crystal and a bulk region of the supercell containing the defect, as described in [148].

Formation energies of charged defects or impurities have to be corrected for the effects of using finite-size supercells. The defect–defect interactions in neighboring supercells converge slowly with the supercell size and are proportional to q^2 where q is the charge state of the defect. It has become clear that the frequently employed Makov–Payne method [151] often significantly overestimates the correction [148, 152], to the point of producing results that are less accurate than the uncorrected numbers. In principle

one can calculate the formation energy of a charged defect for increasing supercell sizes and extrapolate the results to $1/L \rightarrow \infty$, where L is the distance between defects in neighboring supercells. This procedure is computationally expensive and a more rigorous approach (as described in a recent publication [153]) is desirable.

4.2. Defect transition levels

Defects are often electrically active and introduce levels in the band gap of the semiconductor, which involve transitions between different charge states of the same defect [31, 32]. These transition levels are observable quantities that can be derived directly from the calculated formation energies. The transition levels are not to be confused with the Kohn–Sham states that result from band-structure calculations. The transition level $\varepsilon(q/q')$ is defined as the Fermi-level position for which the formation energies of charge states q and q' are equal. $\varepsilon(q/q')$ can be obtained from

$$\varepsilon(q/q') = [E^{\mathrm{f}}(D^q; E_{\mathrm{F}} = 0) - E^{\mathrm{f}}(D^{q'}; E_{\mathrm{F}} = 0)]/(q' - q), \tag{4}$$

where $E^{\mathrm{f}}(D^q; E_{\mathrm{F}} = 0)$ is the formation energy of the defect D in the charge state q when the Fermi level is at the valence-band maximum ($E_{\mathrm{F}} = 0$). The experimental significance of the transition level is that for Fermi-level positions below $\varepsilon(q/q')$ charge state q is stable, while for Fermi-level positions above $\varepsilon(q/q')$ charge state q' is stable. Transition levels can be observed in experiments where the final charge state can fully relax to its equilibrium configuration after the transition, such as in deep-level transient spectroscopy (DLTS) [31, 154]. These thermodynamic transition levels correspond to thermal ionization energies. Conventionally, if a defect transition level is positioned such that the defect is likely to be thermally ionized at room temperature (or at device operating temperatures), this transition level is called a shallow level; if it is unlikely to be ionized at room temperature, it is called a deep level. Note that shallow centers may occur in two cases: first, if the transition level in the band gap is close to one of the band edges (valence-band maximum (VBM) for an acceptor, conduction-band minimum (CBM) for a donor); second, if the transition level is actually a *resonance* in either the conduction or valence band. In that case, the defect necessarily becomes ionized, because an electron (or hole) can find a lower-energy state by transferring to the CBM (VBM). This carrier can still be coulombically attracted to the ionized defect center, being bound to it in a 'hydrogenic effective-mass state'. This second case coincides with what is normally considered to be a 'shallow center' (and is probably the more common scenario). Note that in this case the hydrogenic effective-mass levels that are experimentally determined are *not* directly related to the calculated transition level, which is a resonance above (below) the CBM (VBM).

Note that the transition levels $\varepsilon(q/q')$ are not necessarily the transitions observed in optical spectroscopy experiments. The latter can also be calculated from formation energies where the final and initial charge states correspond to the same defect geometry. In the cases where lattice relaxations of a defect strongly vary from one charge state to another, the thermodynamic transition levels will significantly differ from the optical transition level, which can be obtained by calculating the appropriate configuration coordinate diagrams.

4.3. Migration barriers and diffusion activation energies

In addition to knowing their electronic properties and formation energies, it is also important to know how native point defects migrate in the crystal lattice. Knowledge of migration of point defects greatly contributes to the understanding of their incorporation during growth and processing, and it is essential for modeling self-diffusion and impurity diffusion, which is nearly always mediated by native defects. Information about atomic diffusion or migration of point defects in ZnO is currently limited. Neumann has summarized the experimental results for self-diffusion in ZnO up to 1981 [155]. Activation energies of zinc self-diffusion were reported to be in a range from 1.9 to 3.3 eV, while activation energies for oxygen self-diffusion were reported to span a much wider range, from 1.5 to 7.5 eV. Interpreting these results or using them in a predictive manner is not straightforward. The activation energy for self-diffusion or impurity diffusion (Q) is the sum of the formation energy of the defect that mediates the diffusion process and its migration energy barrier [156]:

$$Q = E^{\mathrm{f}} + E_{\mathrm{b}}. \tag{5}$$

The migration energy barrier E_{b} is a well-defined quantity, given by the energy difference between the equilibrium configuration and the configuration at the saddle point along the migration path, and can be obtained with good accuracy from density-functional calculations [148, 157–159]. The first term in the activation energy, however, namely the defect formation energy (E^{f}), strongly depends on the experimental conditions, such as the position of the Fermi level, and the Zn or O chemical potentials (μ_{Zn} and μ_{O}) in the case of ZnO. Considering that ZnO has a wide band gap of ~3.4 eV, and that μ_{Zn} and μ_{O} can vary in a range of ~3.6 eV given by the ZnO formation enthalpy of ZnO, these parameters can cause large changes in the formation energy. Moreover, it is not straightforward to assess the environmental conditions that affect formation energies and hence the diffusion activation energies in a given experiment. This explains the wide spread in the experimentally determined activation energies, and makes it difficult to extract values from experiment.

In addition to self-diffusion measurements, investigating the behavior of point defects through annealing can also provide valuable information about their migration [46, 111, 160–163]. The point defects in such experiments are often deliberately introduced into the material through non-equilibrium processes such as electron irradiation or ion implantation. Once point defects are introduced, they can be identified by their optical, electronic or magnetic responses (signatures). These signatures are then monitored as a function of annealing temperature. Changes in defect signatures at a given annealing temperature indicate that the relevant defects have become mobile. In principle one can perform a systematic series of annealing experiments at different temperatures,

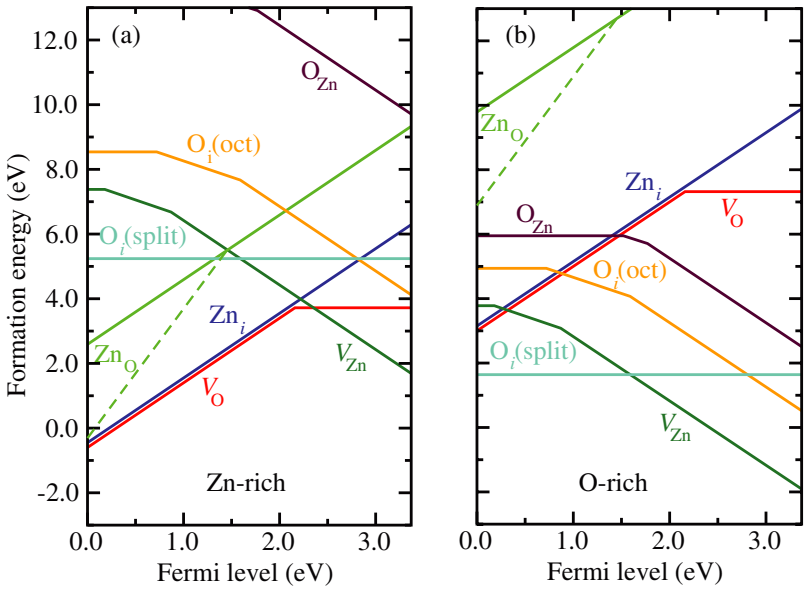

Figure 3. Formation energies as a function of Fermi-level position for native point defects in ZnO for (*a*) Zn-rich and (*b*) O-rich conditions. The zero of Fermi level corresponds to the valence-band maximum. Only segments corresponding to the lowest energy charge states are shown. The slope of these segments indicates the charge state. Kinks in the curves indicate transitions between different charge states. (From Janotti A and Van de Walle C G 2007 *Phys. Rev.* B **76** 165202. With permission.)

extract a time constant for the decay of the signal at each temperature, and then perform an Arrhenius analysis. In the absence of such elaborate studies, estimates for activation energies can still be obtained by performing simpler estimates based on transition state theory [164]. All of this assumes, of course, that the observed changes in defect signatures are solely related to defect migration and do not involve any other processes such as formation of complexes, etc.

4.4. General results of DFT calculations for native defects in ZnO

Density-functional calculations for native defects in ZnO have been reported by several groups [14–29]. However, the fact that the band gap of ZnO is severely underestimated by the commonly used local-density approximation (LDA) or generalized-gradient approximation (GGA) functionals makes the interpretation of the calculations very difficult. Defects often induce occupied states in the band gap. These states have a certain ratio of conduction- versus valence-band character and, therefore, their positions with respect to the VBM can be underestimated by a significant amount. This uncertainty affects the prediction of transition levels and formation energies, leading to potentially large errors, especially in the case of wide-band-gap semiconductors such as ZnO.

Different approaches to overcome the DFT–LDA or GGA deficiencies in predicting band gaps have been employed in the investigation of point defects in ZnO. These include self-interaction corrections, the LDA+U, and the B3LYP and HSE hybrid functionals [17, 20, 22, 24, 25, 27–29]. Although uncertainties still exist in the numerical values of formation energies, important qualitative conclusions can be extracted.

Most of the calculations agree that oxygen vacancies and zinc vacancies are the lowest energy defects, followed by the Zn interstitial and the Zn_O antisite. Oxygen interstitials and O_{Zn} antisites were found to be high in energy. The defects that are favored under Zn-rich conditions (V_O, Zn_i and Zn_O) all act as donors, while those that are favored under O-rich conditions (V_{Zn}, O_i and O_{Zn}) all act as acceptors. In figure 3 we show the calculated defect formation energies as a function of Fermi level from [27]. These results were obtained using an extrapolation scheme based on LDA and LDA+U calculations, as discussed in [20, 22, 27].

Despite the qualitative similarities, it is important to discuss the differences between the results given by the various approaches employed to calculate transition levels and formation energies of native defects in ZnO. Calculations that are based purely on LDA or GGA functionals carry a large uncertainty in the transition levels and formation energies due to the underestimation of the band gap of ZnO by ∼75%. In these cases, transition levels related to defects that induce (single-particle) states in the band gap can be significantly underestimated. When these single-particle states are occupied with electrons, the formation energies for the relevant charge states will be underestimated as well.

In an attempt to overcome this issue, Zhang *et al* included empirical corrections to the bare DFT–LDA results [17]. As a main result, they have found that V_O has a high formation energy in n-type ZnO, with the $\varepsilon(2+/0)$ transition level located in the upper part of the band gap. Lany and Zunger [24] performed LDA+U calculations for perfect ZnO, which partially correct the band gap, and used these results to correct the position of the VBM in ZnO. Otherwise, the results were based on LDA and a rigid shift of the conduction-band minimum (CBM) to correct the band gap, while leaving the positions of deep levels unchanged. Lany and Zunger obtained the V_O $\varepsilon(2+/0)$ transition level at ∼1.6 eV above the VBM.

Using LDA+U, Paudel and Lambrecht concluded that the V_O $\varepsilon(2+/0)$ transition level is located near the VBM [28].

Their scheme includes an application of U to the Zn s states that dominate the character of the conduction band, in addition to applying U to the Zn d states. This seems to go against the nature of the LDA+U correction, which is intended to correct the energies of localized states that are underbound in LDA. While the semicore Zn d states are indeed quite localized, the Zn s states that make up the conduction band are clearly delocalized extended states. Since the V_O-related state in the gap has a large contribution from Zn s states, the application of U to Zn s states will also affect the position of the V_O related state in a way that is, in our opinion, unphysical.

Janotti *et al* observed that LDA+U affects both valence and conduction bands of ZnO [165], and that the single-particle defect states are corrected according to their valence- versus conduction-band character. Because LDA+U only partially corrects the band gap, an extrapolation scheme based on the LDA and LDA+U calculations was then employed to obtain gap-corrected transition levels and formation energies that can be quantitatively compared with experimental results [20, 22, 27].

Using the B3LYP hybrid functional, Patterson carried out calculations of V_O in ZnO [25]. The B3LYP results for the electronic structure of V_O in ZnO are consistent with those obtained by Janotti and Van de Walle [20, 22, 27]. However, Patterson's interpretation of the transition levels based on the results for the single-particle states is not correct. The position of the transition levels cannot be directly extracted from the position of the single-particle states. Transition levels must be calculated based on differences in formation energies (as explained in section 2). This is particularly important for defects which exhibit very different lattice relaxations in different charge states, which as we will see is the case for V_O in ZnO.

Oba *et al* recently performed calculations for point defects in ZnO using the HSE hybrid functional tuned to reproduce the experimental value of the band gap [29]. The calculated single-particle band structure for the oxygen vacancy using the HSE approach is shown in figure 4, indicating that V_O in the neutral charge state induces a doubly occupied single-particle state at ∼1 eV above the VBM [29]. The position of the transition levels in the HSE is in excellent agreement with the results of Janotti and Van de Walle [27]. However, the formation energy of V_O is relatively low, indicating that these defects would be present in significant concentrations in ZnO under extreme Zn-rich conditions. This seems inconsistent with the results of experiments on high-quality ZnO single crystals, in which the electron paramagnetic resonance (EPR) signals identifying oxygen vacancies are not present in as-received crystals [47]. Oxygen vacancies have been observed only after irradiating the samples with high-energy electrons, and are not detected in the as-received samples. It is possible, of course, that these samples were grown or annealed under conditions where the oxygen chemical potential is sufficiently high to suppress oxygen vacancy formation. Indeed, the formation energy of native defects can vary over a wide range (given by the formation enthalpy of ZnO, i.e. 3.6 eV) as a function of the chemical potentials.

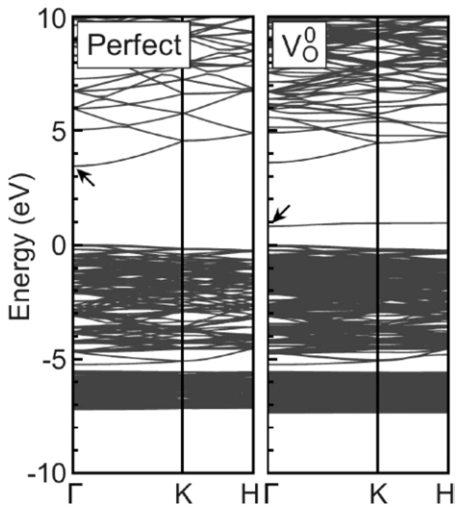

Figure 4. Band structure for the ZnO perfect crystal and for the oxygen vacancy (V_O) in the neutral charge state obtained using the HSE hybrid functional. The zero in energy corresponds to the valence-band maximum in the perfect crystal. (From [29]. With permission.)

4.5. Oxygen vacancies

The oxygen vacancy is the most mentioned defect in the ZnO literature; it is therefore worthwhile devoting special attention to this defect. The oxygen vacancy has most frequently been invoked as the source of unintentional n-type conductivity, but recent research indicates that this assignment cannot be correct. Although the oxygen vacancy has the lowest formation energy among the defects that behave as donors, as shown in figure 3, the density-functional calculations indicate that V_O is a very deep rather than a shallow donor and, consequently, cannot contribute to n-type conductivity [20, 22, 27]. Although the calculations reported in the literature differ on the values for transition levels and formation energies due to the different approaches to correct the band gap, they unanimously agree that V_O is a deep donor [17, 20, 22, 24, 25, 27–29]. According to figure 3, the $\varepsilon(2+/0)$ transition level is located at ∼1 eV below the CBM, i.e. V_O is stable in the neutral charge state in n-type ZnO. The oxygen vacancy is a 'negative-U' center, meaning that $\varepsilon(2+/+)$ lies above $\varepsilon(+/0)$ in the band gap [20, 22, 27]. As the Fermi level moves upward, the charge-state transition is thus directly from the +2 to the neutral charge state.

It should be noted that, while V_O cannot contribute to the n-type conductivity in ZnO because it assumes the neutral charge state when the Fermi level is near the CBM, it can be a relevant source of compensation in p-type ZnO. In this case, V_O assumes the +2 charge state when the Fermi level is near the VBM and has relatively low formation energies as shown in figure 3. In order to avoid incorporation of V_O^{2+} and, hence, avoid compensation in p-type ZnO, it is necessary that during growth or annealing the O chemical potential approaches O-rich conditions and/or the Fermi level is kept away from the VBM, in which cases the formation energy of V_O^{2+} increases.

One can understand the electronic structure of an oxygen vacancy in ZnO based on a simple model within molecular

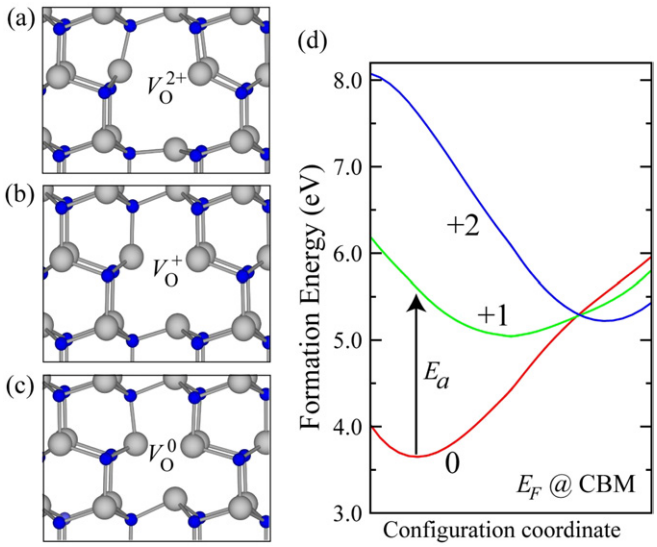

Figure 5. Ball and stick model of the local atomic relaxations around the oxygen vacancy in the (*a*) neutral, (*b*) +1 charge state and (*c*) +2 charge states. In the neutral charge state, the four Zn nearest neighbors are displaced *inward* by 12% of the equilibrium Zn–O bond length. In the +1 charge state, the four Zn nearest neighbors are displaced *outward* by 3%, and for the +2 charge state the displacements are *outward* by 23%. (*d*) Calculated configuration coordinate diagram for V_O^0, V_O^+ and V_O^{2+} for Fermi energy (E_F) at the conduction-band minimum (CBM). The arrow represents an optical transition from V_O^0 to V_O^+ in n-type ZnO. (Based on [20]. With permission.)

orbital theory that involves the four Zn dangling bonds (sp^3 hybrids) and two electrons. These Zn dangling bonds combine into a fully symmetric a_1 state located in the band gap and three almost degenerate states in the conduction band. In the neutral charge state of the oxygen vacancy, the a_1 state is doubly occupied and the three states in the conduction band are empty. From this simple picture, it is easy to see that oxygen vacancy in ZnO can, in principle, exist in three charge states: neutral, +1 and +2, in which the a_1 state is, respectively, doubly occupied, singly occupied and empty.

The occupation of the a_1 state is directly related to the local lattice relaxation around the oxygen vacancy [20, 22, 27]. In the neutral charge state, the four Zn atoms strongly relax inward (toward the vacancy) by 12% of the equilibrium Zn–O bond length; in the +1 charge state they slightly relax outward by 3%; and in the +2 charge state, the four Zn atoms strongly relax outward by 23% as shown in figure 5. These large differences in relaxations significantly reduce the formation energies of V_O^{2+} and V_O^0 relative to V_O^+, leading to a negative-U behavior in which V_O^+ is never thermodynamically stable. This is an important finding, because it is V_O^+, with its unpaired electron, that is detectable by magnetic resonance techniques. An EPR (electron paramagnetic resonance) signal associated with V_O should thus not be observed under thermodynamically stable conditions. It is, of course, possible to create oxygen vacancies in the +1 charge state in a metastable manner, for instance by excitation with light. Once generated, V_O^+ does not immediately decay into the +2 or neutral charge state because of energetic barriers associated with the large difference in lattice relaxations that occur around the oxygen vacancy for the different charge states [20].

As shown in the calculated configuration diagram in figure 5(*d*) [20], the thermal barrier to escape out of the +1 charge state is 0.3 eV, sufficient to maintain an observable concentration of V_O^+ during excitation and cause persistent photoconductivity at low temperatures, but clearly too low to allow for persistent photoconductivity at room temperature as suggested in [24]. Therefore, it is possible to detect EPR signals due to V_O^+ upon photoexcitation at low enough temperatures, but if the excitation is removed and the temperature is raised, these signals decay.

4.5.1. Experimental identification of oxygen vacancies in ZnO. Most of the experimental investigations of oxygen vacancies in ZnO to date have relied on electron paramagnetic resonance (EPR) measurements [38, 40, 47, 166–183, 188]. Many of these studies fall into two categories, depending on the value of the g-factor: one set of reports associates oxygen vacancies with a g value of ∼1.96 [38, 40, 169–181], the other with $g \sim 1.99$ [47, 166–168, 182, 183, 188] (a table containing an overview of these results was included in [27]). There is, however, overwhelming evidence that oxygen vacancies are actually associated with the $g \sim 1.99$ line. For example, the $g \sim 1.99$ signal has only been observed after irradiation of the samples, indicating that it is related to a native defect (and consistent with the result that V_O has a high formation energy and is thus unlikely to occur in as-grown n-type material). Also, it has been found that illumination is necessary to observe the center [47, 166–168], consistent with the results of density-functional calculations in [20] which indicate that excitation is required in order to generate the paramagnetic +1 charge state.

In addition, hyperfine interactions with the ^{67}Zn neighbors of the vacancy were observed for the $g \sim 1.99$ line [166, 167], whereas no hyperfine interactions have been reported for the $g \sim 1.96$ line. The latter is likely to be associated with electrons in the conduction band or in a donor band, as originally proposed by Müller and Schneider [171] and most recently confirmed by Garces *et al* [181]. The erroneous association of the $g \sim 1.96$ line with V_O is probably related to the prevailing hypothesis that oxygen vacancies were the donors responsible for the unintentional n-type conductivity in ZnO. Note that Sancier in 1970 also favored assigning the $g \sim 1.96$ line to electrons in the conduction band [174], and Neumann in 1981 observed that doping with Al, Ga or In increases the intensity of the $g \sim 1.96$ signal. The $g \sim 1.96$ signal has also been reported to be enhanced under UV illumination [169, 171–176, 178]. UV light can indeed promote electrons into the conduction-band states in ZnO, consistent with the $g \sim 1.96$ line corresponding to electrons in delocalized states.

Leiter *et al* have performed photoluminescence and optically detected electron paramagnetic resonance (ODEPR) measurements in as-grown (i.e. unirradiated) single-crystal ZnO [185, 186]. They observed a spin-triplet signal ($S = 1$) with $g_{||} = 1.984$, $g_\perp = 2.025$ which they attributed to oxygen vacancies based on analogies with anion vacancies in other oxides. This assignment disagrees with the experiments cited

above, which relate oxygen vacancies with g values of 1.99. In addition, the signal observed by Leiter *et al* is present already in the as-grown material, unlike the observations of the $g \sim 1.99$ signal, which all required irradiation. More recently, based on PL and ODEPR measurements, Hofmann *et al* [187] suggested a correlation between the often-observed broad green emission centered at 2.45 eV and a donor level 530 meV below the conduction band which was assigned to the $\varepsilon(2+/0)$ of the oxygen vacancy. The 2.45 eV emission band would be related to a transition from a triplet ($S = 1$) excited state to the singlet ground state of the neutral oxygen vacancy. These intriguing suggestions have not yet been verified. Based on the results of DLTS measurements combined with optical excitation, Hofmann *et al* [187] also proposed that the $\varepsilon(2+/+)$ level is located at 140 meV below the conduction band, thus indicating the negative-U nature of the oxygen vacancy in ZnO.

Vlasenko and Watkins have also carried out ODEPR measurements in high-quality ZnO single crystals [47], whose results are in good agreement with the first-principles results shown in figures 3 and 5. They noted that the EPR signals related to V_O could be detected only after high-energy electron irradiation, indicating that oxygen vacancies are not present in the as-grown (or as-received) ZnO single crystals. Recent experiments by Evans *et al* also confirm that the EPR signals related to oxygen vacancies are not detectable in as-grown ZnO single crystals, but only after irradiation [188]. Moreover, Vlasenko and Watkins have reported that V_O^+ can be observed only upon excitation with photon energies above ~2 eV [47], while Evans *et al* report a threshold excitation energy of 2.1 eV [188], in good agreement with earlier measurements which reported that the peak for optical transition $V_O^0 \rightarrow V_O^+$ occurs at 2.3 eV [167, 189]. These results confirm that V_O^+ is not thermodynamically stable as discussed in [20, 27]. Moreover, the excitation energy is in good agreement with the optical transitions extracted from the calculated configuration coordinated diagram from [20], reproduced in figure 5(*d*).

In addition to EPR studies, there exist a few investigations of oxygen vacancies in ZnO using positron annihilation spectroscopy measurements [111, 190]. In these reports, the samples were electron irradiated and had a Fermi level 0.2 eV below the CBM after irradiation. The dominant compensating defect was found to be the zinc vacancy; however, the measurements also produced evidence for the presence of a neutral defect, which was proposed to be the neutral oxygen vacancy. These observations are fully consistent with the results of density-functional calculations reported in [20, 22, 27], both regarding the absence (below the detection limit) of oxygen vacancies in the as-grown material and V_O being present in the neutral charge state when E_F is 0.2 eV below the CBM.

4.5.2. Migration energy barrier for V_O. In the migration of oxygen vacancy, a nearest-neighbor oxygen atom in the oxygen lattice jumps into the original vacant site leaving a vacancy behind. The migration energy barrier for this process can be obtained by calculating the total energy at various intermediate configurations when moving a neighboring oxygen atom from its nominal lattice site along a path toward the vacancy. The oxygen vacancy has twelve next-nearest-neighbor oxygen atoms—six are located in the same basal plane as the vacancy, accounting for vacancy migration perpendicular to the c axis, and the other six neighbors are located in basal planes above and below the basal plane of the oxygen vacancy, accounting for vacancy migration both parallel and perpendicular to the c axis. The calculations reported in [27] indicate that the migration of oxygen vacancies is isotropic, i.e. migration barriers involving oxygen atoms from the basal plane of the vacancy and from planes above or below the basal plane of the vacancy have the same value. However, as reported in [27], the migration barrier does depend on the charge state of the oxygen vacancy. The calculated migration barrier for V_O^0 is 2.4 eV and for V_O^{2+} is 1.7 eV. The former is relevant in n-type material whereas the latter is relevant in semi-insulating (E_F below 2.2 eV) or p-type materials. These energy barriers indicate that V_O^0 will become mobile above temperatures 900 K and V_O^{2+} will become mobile at temperatures above 650 K [27].

First-principles calculations of migration barriers for V_O in ZnO have also been reported by Erhart and Albe [23]. However, they found differences as large as 0.7 eV between migration barriers involving oxygen atoms from the basal plane of the vacancy and from planes above or below the basal plane of the vacancy. Such large anisotropies in the migration barriers are quite unexpected—the local geometry around V_O is almost tetrahedrally symmetric—and are probably an artifact of using rather small supercells (32 atoms) in the calculations of migration barriers [23]. In fact, calculated migration barriers for a nitrogen vacancy in GaN using 32- and 96-atom supercells differ by as much as 0.6 eV, due to the large relaxations surrounding the vacancy, which are not properly described in the 32-atom supercell [158].

4.6. Zinc vacancies

The electronic structure of zinc vacancies in ZnO can also be understood using a simple model within molecular orbital theory. The removal of a Zn atom from the ZnO lattice results in four O dangling bonds and a total of six electrons; these four O dangling bonds combine into a doubly occupied symmetric a_1 state located deep in the valence band, and three almost degenerate states in the band gap, close to the VBM. These three states are partially occupied by a total of four electrons and, therefore, can accept up to two additional electrons, explaining the acceptor behavior of V_{Zn} in ZnO. Because the formation energy of acceptor-type defects decreases with increasing Fermi level, V_{Zn} can more easily form in n-type materials. Zinc vacancies have very high formation energies in p-type ZnO as shown in figure 3, and therefore their concentration should be negligibly low. In n-type ZnO, on the other hand, V_{Zn} have the lowest formation energy among the native point defects, indicating that V_{Zn}^{2-} can occur in modest concentrations in n-type ZnO, acting as a compensating center. In fact, V_{Zn} have been identified as the dominant compensating center in n-type ZnO by positron annihilation measurements [111, 127]. They are also more favorable in oxygen-rich conditions as shown in figure 3.

According to the calculations reported in [27], V_{Zn} are deep acceptors with transition levels $\varepsilon(0/-) = 0.18$ eV and

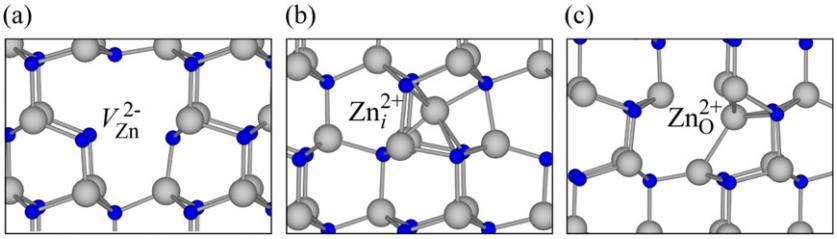

Figure 6. Calculated local atomic geometry for (*a*) the Zn vacancy in the −2 charge state (V_{Zn}^{2-}; (*b*) the zinc interstitial in the +2 charge state (Zn_i^{2+}). (*c*) Calculated local atomic geometry of the zinc antisite in the +2 charge state (Zn_O^{2+}). (From Janotti A and Van de Walle C G 2007 *Phys. Rev.* B **76** 165202. With permission.) (Color online.)

$\varepsilon(-/2-) = 0.87$ eV. Due to their very high formation energy in p-type materials, V_{Zn} are unlikely to play any role in p-type conductivity. Recently it has been proposed that a complex formed by a column-V element (As or Sb) on a substitutional Zn site surrounded by two zinc vacancies would be responsible for p-type conductivity [70]. From figure 3, the formation energy of a single V_{Zn} is 3.7 eV under p-type conditions (E_F at the VBM), in the most favorable extreme O-rich limit. The formation energy of an As_{Zn}–$2V_{Zn}$ complex would then amount to 2×3.7 eV plus the formation energy of substitutional As_{Zn} (or Sb_{Zn}) minus the binding energy of the complex. Assuming that the formation energy of As_{Zn} (or Sb_{Zn}) is about 1 eV (a very low value considering the chemical mismatch between Zn and As), a binding energy of more than 6 eV would be required to stabilize these complexes at equilibrium conditions. This value is too large to be attainable, and indeed much larger than the value calculated in [70]. Therefore we feel that these complexes cannot be responsible for p-type conductivity.

The oxygen atoms around V_{Zn} exhibit large outward breathing relaxations of about 10% with respect to the equilibrium Zn–O bond length (figure 6). Similar relaxations are observed for the three possible charge state configurations V_{Zn}^{0}, V_{Zn}^{-} and V_{Zn}^{2-}. This indicates that the overlap between the oxygen 2p orbitals surrounding the zinc vacancy is too small to result in significant chemical bonding between the oxygen atoms. Indeed, the calculated O–O distances are about 3.5 Å, much larger than the sum of the covalent radii (1.46 Å).

4.6.1. EPR studies of zinc vacancies. Several EPR observations of zinc vacancies in ZnO have been reported. Taylor *et al* [182] reported EPR signals with g factors in the range 2.0018–2.056 in irradiated single crystals. It was proposed that a subset of these lines would be due to Zn vacancies. Galland and Herve [191] observed lines with g factors between 2.0024 and 2.0165 in irradiated single crystals, also attributing them to Zn vacancies. In addition, it has recently been suggested that the triplet $S = 1$ signal (with $g_{||} = 1.984$, $g_{\perp} = 2.025$) observed by Leiter *et al* [185, 186] may be related to spin-dependent recombination involving a $S = 1/2$ V_{Zn}^{-} exchange-coupled to a $S = 1/2$ effective-mass donor [27].

4.6.2. Migration of V_{Zn}. In analogy to the case of V_O, migration of V_{Zn} occurs when a nearest-neighbor Zn atom moves into the vacant site leaving a vacancy behind. The migration of V_{Zn}^{2-} in ZnO has been predicted to be isotropic with a migration barrier of 1.4 eV [27]. This value of migration barrier indicates that V_{Zn} will become mobile at temperatures above 540 K, in good agreement with positron annihilation measurements which indicate that V_{Zn} can be annealed out at temperatures between 500 and 600 K [111].

4.6.3. Zinc vacancies and the origin of the green luminescence. ZnO often exhibits a broad and weak green luminescence, centered between 2.4 and 2.5 eV [192–195]. This green luminescence has been observed in samples prepared with a variety of growth techniques, and it is important to point out that *there may not be a single source for this luminescence*. For instance, the presence of Cu impurities has been suggested as a potential cause [196, 197]; nevertheless, not all ZnO samples for which the green luminescence has been observed contain Cu. Native defects have also been suggested as a potential source. Reynolds *et al* [193, 194] and Kohan *et al* [15] have suggested that the Zn vacancy can give rise to green luminescence. Indeed, the calculated transition level between the −1 and −2 charge states occurs at 0.9 eV above the VBM [27], and hence a transition between the conduction band (or a shallow donor) and the V_{Zn} acceptor level would give rise to luminescence around 2.5 eV, in good agreement with the observed transition energy. In addition to the agreement with the observed emission energy, the Zn vacancy is also a likely candidate because it is an acceptor-type defect: acceptor defects are more likely to occur in n-type material, and most ZnO samples to date have exhibited unintentional n-type conductivity. Note that this proposed explanation for the green luminescence is similar to the proposal that gallium vacancies are the source of the yellow luminescence in GaN [198].

Other explanations have been proposed for the green luminescence. Several groups have suggested that oxygen vacancies are the source of green luminescence [17, 179, 185, 199–201]. Vanheusden *et al* reported a correlation between the intensity of the green luminescence and the concentration of oxygen vacancies [179, 201]. However, their assessment of the presence of oxygen vacancies was based on the observation of a line with $g \sim 1.96$ in EPR measurements; as discussed above this assignment is not correct, undermining the arguments made in [179, 201]. The calculated configuration-coordinate diagrams for V_O reported in [20] also do not show any transitions consistent with green luminescence. Leiter *et al* [185, 186] and Hofmann *et al* [187] associated the emission band around 2.45 with oxygen

vacancies based on optically detected magnetic resonance experiments. As discussed in section 4.5.1, it is unlikely that the $S = 1$ center that they observed represents the oxygen vacancy. Instead, V_{Zn} may be the cause of the observed green luminescence in their experiments. Indeed, a $S = 1$ EPR signal related to V_{Zn} has recently been proposed by Vlasenko and Watkins [46].

Experiments by Sekiguchi *et al* also provide a strong argument in favor of V_{Zn} being the source of green luminescence [202]. They reported a strong passivation of the green luminescence by hydrogen plasma treatment. This observation is consistent with the green luminescence being caused by V_{Zn}, which act as acceptors and can be passivated by hydrogen [14, 203]. In fact, the same passivation effect was observed by Lavrov *et al*, who simultaneously observed an increase in vibrational modes associated with hydrogenated Zn vacancies [203]. Indeed, the passivation of the green luminescence by hydrogen is very plausible if the green luminescence is caused by zinc vacancies, as H atoms passivate V_{Zn} by forming strong O–H chemical bonds [203].

4.7. Zinc interstitials

A Zn interstitial atom (Zn_i) could in principle occupy the tetrahedral site or the octahedral site in the ZnO wurtzite structure. At the tetrahedral site, the Zn_i has one Zn and one O as nearest-neighbor atoms, at a distance of $\sim 0.833 d_0$ (d_0 is the Zn–O bond length along the c axis). At the octahedral site, the Zn_i has three Zn and three O atoms as nearest neighbors at a distance of $\sim 1.07 d_0$. Based on size considerations, it is therefore expected that the Zn interstitial will be more stable at the octahedral site where the geometrical constraints are less severe. Indeed it has been found that the octahedral site is the stable site for Zn_i, whereas the Zn_i at the tetrahedral site is 0.9 eV higher in energy and unstable as it spontaneously relaxes to the octahedral site [22, 27]. Moreover, instead of occupying the ideal octahedral site, the density-functional calculations indicate a large displacement of Zn_i along the c-axis, resulting in an increased Zn_i–Zn distance of $1.22 d_0$ and a decreased Zn_i–O distance of $1.02 d_0$. Similar displacements along the c axis were also observed for Ga_i in GaN [158].

Regarding its electronic structure, Zn_i induces an a_1 state with two electrons above the CBM. These two electrons are transferred to conduction-band states, stabilizing the +2 charge state (Zn_i^{2+}). Hence, Zn_i will always donate electrons to the conduction band, thus acting as a shallow donor [27]. These electrons can, of course, be bound to the defect center in hydrogenic effective-mass states, so that effectively the observed defect levels would be the effective-mass levels below the CBM. From figure 3, we note that Zn_i has a high formation energy in n-type samples, in which the Fermi level lies near the CBM, and even under extreme Zn-rich conditions. Zn interstitials are therefore unlikely to be responsible for unintentional n-type conductivity, since they will be present in very low concentrations in n-type ZnO. On the other hand, the formation energy of Zn_i^{2+} decreases rapidly when the Fermi level decreases toward the VBM, making Zn interstitials a potential source of compensation in p-type ZnO.

It has been suggested, however, that Zn_i can be introduced in n-type ZnO under non-equilibrium conditions. Thomas [36] reported the presence of shallow donors when ZnO crystals were heated in Zn vapor followed by a rapid quench, and Hutson [35] observed the appearance of a shallow donor with ionization energy of 51 meV in Hall measurements. Look *et al* conducted high-energy electron irradiation experiments and identified a shallow donor with an ionization energy of 30 meV [112]. Based on the much higher production rate of this defect for the Zn $(0\,0\,0\,1)$ face than for the O $(0\,0\,0\,\bar{1})$ face, the donor was suggested to be related to a Zn-sublattice defect, either the Zn interstitial itself or a Zn-interstitial-related complex.

Recent density-functional calculations revealed that Zn interstitial migrates through a kick-out mechanism with a very low migration barrier of 0.57 eV [27]. This result is in very good agreement with early experiments that reported a migration barrier of 0.55 eV for Zn interstitials in ZnO crystals heated in Zn vapor followed by rapid quenching [36]. This low migration barrier implies that Zn_i is highly mobile even at room temperature. Due to its high formation energy and high mobility, it is therefore unlikely that *isolated* Zn_i was the observed species in the experiments mentioned above. Alternatively, a Zn-interstitial-related complex could be responsible for the observations. Look *et al* suggested that Zn_i–N_O would form under N ambient conditions and contribute to the unintentional n-type conductivity [204].

Due to its high formation energy, it is also unlikely that Zn_i mediates Zn self-diffusion in the ZnO samples that have been studied to date, which are invariably n-type. Instead, it is more likely that Zn self-diffusion is mediated by Zn vacancies which have a higher migration barrier of 1.40 eV, but have a much lower formation energy in n-type ZnO [27]. Indeed, Tomlins *et al* reported an activation energy of 3.86 eV for self-diffusion of Zn in ZnO, and appropriately suggested that Zn self-diffusion is controlled by a vacancy mechanism [205].

4.8. Zinc antisites, oxygen interstitials and oxygen antisites

The other native defects (zinc antisites, oxygen interstitials and oxygen antisites) have higher formation energies and, therefore, are not expected to play a role in ZnO under near-equilibrium conditions. The Zn_O antisite is a double donor in n-type ZnO, but its high formation energy indicates that it is an unlikely source of unintentional n-type conductivity. Recent first-principles calculations revealed a large off-site displacement of the Zn atom by more than 1 Å from the substitutional lattice site toward two next-nearest-neighbor oxygen atoms along the $[1\,0\,\bar{1}\,0]$ direction, as shown in figure 6(*c*) [22, 27]. The resulting Zn_O–O interatomic distances are only 8% larger than the equilibrium Zn–O bond length. At this equilibrium configuration we find three Zn_O–Zn distances of ~ 2.4 Å and one Zn_O–Zn distance of ~ 2.8 Å. One can think of this low-symmetry configuration of Zn_O as a complex of a Zn interstitial and an O vacancy [27].

The migration of Zn_O would involve splitting the defect into its Zn_i and V_O constituents, raising the question of whether these constituents would remain bound and move in concert, or

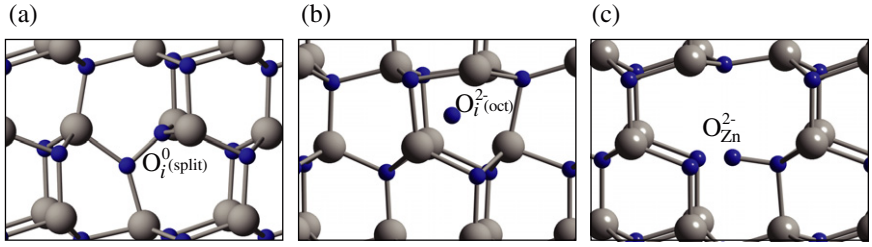

Figure 7. Local atomic geometries for (*a*) oxygen interstitial in the split-interstitial configuration O_i^0(split); (*b*) oxygen interstitial in the octahedral configuration O_i^{2-}(oct); (*c*) oxygen antisite O_{Zn}^{2-}, showing a large displacement off the substitutional site. (From Janotti A and Van de Walle C G 2007 *Phys. Rev.* B **76** 165202. With permission.) (Color online.)

would move independently. The estimated dissociation barrier of Zn_O^{2+} into Zn_i^{2+} and V_O^0 is 1.3 eV in n-type ZnO. Zn_O^{2+} is therefore expected to be stable at temperatures of up to ∼500 K [27]. This suggests that Zn_O^{2+} may cause n-type conductivity if intentionally introduced under non-equilibrium conditions such as high-energy electron irradiation [112]. Indeed, recent experiments on electron-irradiated ZnO reported that the donor defects created by irradiation anneal out at 400 °C [204], consistent with the stability of Zn_O^{2+}.

The oxygen interstitial (O_i) can, in principle, occupy the octahedral or tetrahedral interstitial sites, or form split interstitials. Density-functional calculations indicate that the O_i at the tetrahedral site is unstable and spontaneously relaxes into a split-interstitial configuration in which the extra O atom shares a lattice site with one of the nearest-neighbor O atoms [27]. In the O_i(split) configuration, the O–O distance is 1.46 Å (figure 7(*a*)), suggesting the formation of an O–O chemical bond, with two almost degenerate and completely filled states in the band gap that resemble the antibonding ppπ^* state in a molecular orbital description of the isolated O_2 molecule. In contrast to the isolated molecule in which the ppπ^* molecular orbital (MO) is occupied by two electrons with parallel spins (resulting in a triplet $S = 1$ ground state), the ppπ^*-like MO of O_i(split) is completely occupied [27]. This explains the significantly longer O–O bond length in the O_i(split) (1.46 Å) compared with that of the isolated O_2 molecule (1.22 Å). As a result, O_i(split) is electrically inactive, i.e. it is neutral for any Fermi-level position (see figure 3), with the calculated donor transition levels $\varepsilon(+/0)$ and $\varepsilon(+/2+)$ occurring below the VBM [27].

It has been found that O_i(split) has a metastable configuration with formation energy ∼0.2 eV higher than the lowest energy configuration, with O–O bond length 1.51 Å [27]. The existence of these two almost degenerate configurations with very different O–Zn–O bond angles reinforces the picture of the oxygen split interstitial as an O_2 molecule embedded in the ZnO crystal [27, 206]. Erhart and Albe [23] also reported on the O_i(split) and its metastable configuration, and suggested that these defects are electrically active with acceptor transition levels close to the CBM. We believe that the acceptor charge states found by Erhart and Albe are a result of occupying extended bulk states near the CBM, and not defect-induced states.

Oxygen interstitials can also exist as electrically active interstitials occupying the octahedral site, O_i(oct) as shown in figure 7(*b*), introducing states in the lower part of the band gap that can accept two electrons. These states are derived from oxygen *p* orbitals and result in deep acceptor transition levels $\varepsilon(0/-)$ and $\varepsilon(-/2-)$, at 0.72 and 1.59 eV above the VBM. As shown in figure 3, O_i can exist either as electrically inactive O_i^0(split) in semi-insulating and p-type materials, or as deep acceptors at the octahedral interstitial site O_i^{2-}(oct) in n-type materials ($E_F > 2.8$ eV) [27]. Note that for both forms the formation energies are very high (except under extreme O-rich conditions). The calculated migration energy barrier for O_i^0(split) is 0.9 eV, while the migration barrier for O_i^{2-}(oct) through the hexagonal channel along the *c* axis is 1.1 eV. These values indicate that O_i^0(split) will become mobile at temperatures above 340 K and O_i^{2-}(oct) will become mobile at $T > 440$ K [27]. Based on a comparison with the migration energies of V_O (discussed in section 4.5.2), it is suggested that O_i is the species responsible for defect recombination on the oxygen sublattice, especially under n-type conditions.

The oxygen antisite (O_{Zn}) is an acceptor-type defect with very high formation energy, even under the most favorable O-rich conditions as shown in figure 3 [27]. Hence, it is very unlikely that O_{Zn} would be present in equilibrium. However, O_{Zn} could potentially be created under non-equilibrium conditions such as under irradiation or ion implantation. It was found that O on the ideal Zn site is unstable and spontaneously relaxes to an off-site configuration [27]. The O atom is displaced along the $[0\,0\,0\,\bar{1}]$ direction by more than 0.7 Å, and forms a chemical bond with one of the oxygen nearest neighbors as shown in figure 7(*c*). The O–O bond length is 1.46 Å (exactly twice the covalent radius) in the −2 charge state and 1.42 Å in the neutral charge state. The distances between O_{Zn} and the other nearby oxygen atoms are ∼2.0 Å, much larger than twice the oxygen covalent radius of 0.73 Å, thus indicating the absence of bonding. O_{Zn} are deep acceptors with transition levels $\varepsilon(0/-)$ and $\varepsilon(-/2-)$ at 1.52 and 1.77 eV above the VBM [27]. In order to migrate, O_{Zn} would dissociate into V_{Zn} and O_i and, qualitatively, it is expected that its migration barrier will be higher than that of vacancies or interstitials.

It has been suggested by Lin *et al* that the green luminescence in ZnO corresponds to deep levels induced by oxygen antisites [195]. They observed that the intensity of the green emission increases with oxygen partial pressure in thin films annealed at high temperatures, and ruled out V_{Zn} and O_i based on previous density-functional calculations of the defect transition levels. However, O_{Zn} has much higher formation energy than both V_{Zn} and O_i, and is therefore unlikely to

be present in significant concentrations under equilibrium conditions [27]. As discussed in section 4.6.3, V_{Zn} are more likely to be the cause of the observed green luminescence in ZnO, and equally consistent with the results of Lin *et al* [195].

4.9. General remarks about native defects in ZnO

To summarize what is known about native defects, oxygen vacancies are deep donors and cannot explain the observed n-type conductivity in ZnO. However, they can compensate p-type doping. Zinc interstitials are shallow donors, but have high formation energies under n-type conditions; moreover, they are fast diffusers and hence unlikely to be stable as isolated point defects. Zinc antisites are also shallow donors, but they have high formation energies in n-type samples. Zinc antisites show a large off-site displacement and induce a large local lattice relaxation.

Zinc vacancies are deep acceptors and have low formation energies under n-type conditions; they can therefore occur as compensating defects in n-type samples. It is suggested that zinc vacancies are a possible source of the often-observed green luminescence in ZnO. Oxygen interstitials have high formation energies and are not expected to exist in significant concentrations. They can exist as electrically inactive split interstitials or as deep acceptors at the octahedral site in n-type samples. Oxygen antisites have the highest formation energies among the acceptor-type native point defects. They are deep acceptors and also show large off-site displacements, in which the oxygen atom bonds chemically to only one of the oxygen nearest neighbors.

The migration barriers of all the point defects are modest, explaining why radiation damage can be annealed out at relatively low temperatures (with some recovery already taking place below room temperature). Zinc interstitials diffuse through the kick-out mechanism with a rather low migration barrier of 0.57 eV, in agreement with experimental observations, and are responsible for the observed fast recovery of the electrical properties in irradiated ZnO. The migration barrier of oxygen interstitials in the octahedral configuration (relevant for n-type samples) is 1.1 eV, whereas zinc and oxygen vacancies diffuse with somewhat higher migration barriers of 1.4 eV and 2.4 eV, respectively. We note that these low migration barriers imply that most point defects will be highly mobile at the temperatures at which ZnO crystals and epilayers are commonly grown, indicating that such growth processes can be considered to be near equilibrium.

5. Donor impurities

Having established that native point defects cannot explain the unintentional n-type conductivity in ZnO, it is relevant to investigate possible impurities that act as shallow donors and are commonly observed in as-grown ZnO crystals. For the sake of controlling the n-type conductivity in ZnO, it is also relevant to investigate other possible impurities that act as shallow donors and can be used to make n-type ZnO in a stable manner. The properties of dopant impurities in ZnO are summarized in table 1.

Table 1. Properties of dopant impurities in ZnO. n_{max} or p_{max} indicates the highest carrier concentration experimentally observed to date.

Impurity	Character	Ionization energy	n_{max} or p_{max} (cm^{-3})
Al	Donor	120 meV [17]	8×10^{20} [207]
Ga	Donor	—	1.1×10^{20} [208]
		—	3.7×10^{20} [210]
In	Donor	—	[211]
F	Donor	80 meV [17]	5×10^{20} [213]
H	Donor	35 meV [49]	—
Li	Acceptor	—	
Cu	Acceptor	—	
N	Acceptor	100 meV [62]	9×10^{16} [62]

5.1. Boron, aluminum, gallium and indium

The group-III impurities B, Al, Ga and In when substituted on the Zn site act as shallow donors in ZnO. The extra valence electron of these impurities is loosely bound and occupies effective-mass states near the CBM at low temperatures. As the temperature rises this extra electron is excited to the conduction band and is free to move. For Al, Zhang *et al* [17] calculated an ionization energy of 120 meV and a low formation energy, referenced to the elemental Al phase. This formation energy would be raised if equilibration with Al_2O_3 were taken into account. Hu and Gordon [207] obtained carrier concentrations up to 8×10^{20} cm^{-3} with Al doping in chemical vapor deposition of ZnO. For Ga, Ko *et al* [208] found that high Ga doping (up to 10^{20} cm^{-3}) did not degrade the structural quality of the film, but a degradation in the photoluminescence intensity was observed for Ga concentrations exceeding 2.6×10^{19} cm^{-3}. Indium has also been used as a donor in ZnO [211]. An EPR study for In in ZnO was carried out by Block *et al* [209]. Boron has also been used as n-type dopant in ZnO thin films grown by MOCVD [212].

5.2. Fluorine

Fluorine has one more electron than O, and when inserted on the O site it acts as a shallow donor in ZnO. Zhang *et al* [17] calculated an ionization energy of 80 meV and a low formation energy, indicative of easy incorporation into the lattice. Intentional fluorine doping can produce electron concentrations up to 5×10^{20} cm^{-3} [213]. The atomic structure of the fluorine donor is unexpected: for F^+, a large displacement of one of the neighboring Zn atoms is found. This Zn atom moves away from the F atom by 25% of the bond length; the F atom itself moves off-site by 12% of the bond length [16]. The bond between the F atom and one of its Zn neighbors is thus effectively broken. Such large relaxations are usually thought of as giving rise to deep (localized) states, but F in ZnO still behaves as a shallow donor.

5.3. Hydrogen

Recent first-principles calculations have drawn attention to the role of hydrogen in the electronic properties of ZnO [14, 26]. It was found that interstitial hydrogen behaves as a shallow donor: only the positive charge state (H_i^+) is

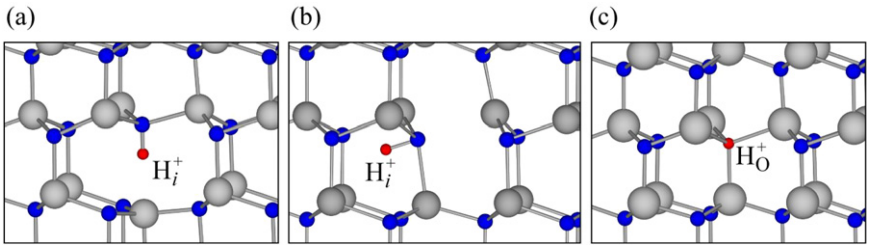

Figure 8. Ball and stick model of the relaxed atomic positions of interstitial hydrogen: (*a*) at the bond-center site parallel to the *c* axis; (*b*) at the antibonding site perpendicular to the *c* axis; (*c*) at the substitutional oxygen site. (Color online.)

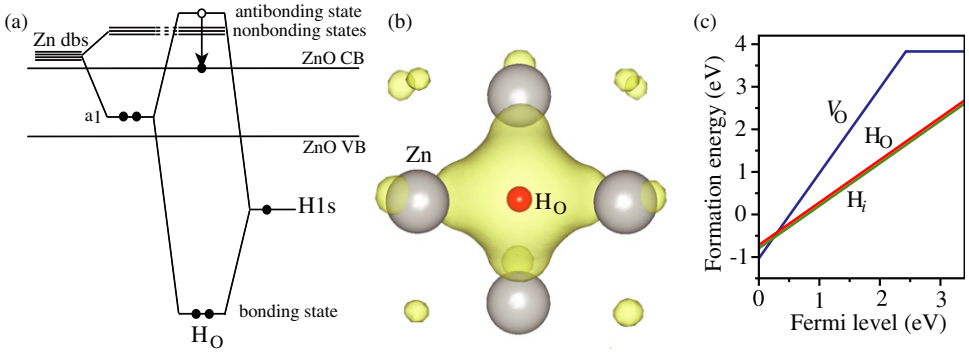

Figure 9. (*a*) Coupling between the H 1s orbital and the Zn 4s dangling bonds (Zn dbs) to form the hydrogen multicenter bond in ZnO. The H 1s orbital combines with the a_1 state and results in a fully symmetric bonding state in the valence band, and an antibonding state in the conduction band. The electron that would occupy this antibonding state is then transferred to the conduction-band minimum, making the substitutional hydrogen H_O a shallow donor. (*b*) Electronic charge density distributions of the lowest energy fully symmetric bonding state of the hydrogen multicenter bond in ZnO. The isosurfaces are at 0.05 electrons Å^{-3}. (*c*) Formation energies as a function of the Fermi-level position for substitutional hydrogen H_O, interstitial hydrogen H_i and the oxygen vacancy V_O in ZnO under Zn-rich conditions. The slopes correspond to the charge of the stable states. H_O and H_i are both stable in the +1 charge state. The oxygen vacancy is stable only in the +2 and 0 charge states. The Fermi level is referenced to the valence-band maximum. (From [26]. With permission.)

thermodynamically stable. An electron can be bound to the H_i^+ center in an extended state, of course, characteristic of a shallow donor. The behavior of hydrogen in ZnO is highly unusual. In almost all semiconductors studied to date, interstitial hydrogen has been found (theoretically as well as experimentally) to act as an *amphoteric* impurity [214–216]: in p-type material, hydrogen incorporates as H_i^+, and in n-type material as H_i^-, always counteracting the prevailing conductivity of the material. This amphoteric behavior precludes hydrogen from acting as a dopant, i.e. from being a *source* of conductivity. In ZnO, however, interstitial hydrogen occurs exclusively in the positive charge state, i.e. it always acts as a donor.

Interstitial hydrogen in ZnO can be located at the bond-center (BC) site, or at the antibonding (AB_O) site, with comparable energies. Interstitial hydrogen in ZnO thus prefers sites where it can strongly bind to an oxygen atom, forming an O–H bond with a length of 0.99–1.01 Å. Large lattice relaxations occur around the hydrogen interstitial; in particular, for the bond-center (BC) configuration the Zn atom moves outward over a distance equal to 40% of the bond length (0.8 Å), to a position slightly beyond the plane of its nearest neighbors as shown in figure 8(*a*). Simultaneously, the O atom moves outward by 11% of the bond length. For the AB_O configuration, the relaxation of both Zn and O amounts to about 20% of the bond length as shown in figure 8(*b*). As we noted in section 5.2, such large relaxations are not unique to hydrogen; they also occur for fluorine in ZnO [16].

In addition to the interstitial positions, it was recently found that hydrogen can also replace oxygen in ZnO (H_O), forming a multicenter bond in which H is equally bonded to the four Zn nearest neighbors [26], as shown in figure 8(*c*). Substitutional hydrogen H_O is also a shallow donor in ZnO, occurring exclusively in the positive charge state H_O^+ [26]. The hydrogen multicenter bond can be understood as a coupling between the fully symmetric state of V_O and the H 1s state, and is located at ∼7 eV below the VBM as shown in figure 9(*a*). The electronic charge density of the hydrogen multicenter bond in ZnO is shown in figure 9(*b*).

The substitutional and interstitial forms of hydrogen have low formation energies in ZnO, as shown in figure 9(*c*), indicating that they can occur in significant concentrations. Hydrogen is obviously not the only possible donor in ZnO, but it is a very attractive candidate for an impurity that can be unintentionally incorporated and can give rise to background n-type conductivity. Hydrogen is either intentionally or unintentionally present in the growth environment of the most used techniques used to produce ZnO. These techniques include vapor-phase transport [76], hydrothermal growth [105], chemical vapor deposition (MOCVD) [217], laser ablation [109] and sputtering [110]. In addition, processing steps such as wet etching or annealing in forming gas can easily introduce hydrogen into the material.

Experimental indications for hydrogen's behavior as a donor in ZnO were already reported in the 1950s [51, 218, 219]. Thomas and Lander [219] observed an increase in n-type conductivity when H was diffused into ZnO. They used the measured conductivity as a function of temperature to derive the solubility of H in ZnO, and found the heat of the reaction $H_2(\text{gas}) \rightarrow 2H^+ + 2e$ to be 3.2 eV or 1.6 eV per hydrogen. This value should correspond (to within small correction terms) to the formation energy of interstitial H^+ in ZnO, and it is in good agreement with the first-principles calculations [14, 26].

An increase in conductivity upon exposure to H_2 was also observed by Baik *et al* [220] and by Kohiki *et al* [221], who introduced hydrogen by proton implantation followed by annealing at 200 °C. Low-energy H implantation of ZnO surfaces also results in formation of electron accumulation layers, consistent with H acting as a donor [222].

The nature of hydrogen as a donor impurity has been microscopically established in recent experiments [48, 49]. Muon spin rotation is a technique similar to electron paramagnetic resonance (EPR), based on muonium, which is a pseudo-isotope of hydrogen. Muonium in ZnO was observed to exhibit all characteristics of a shallow donor (including ionization behavior consistent with a level close to the conduction band, and a delocalized wavefunction), confirming that H acts as a shallow donor [48]. Electron paramagnetic resonance has also resulted in a direct observation of hydrogen in ZnO, with behavior consistent with a shallow donor [49]. Hydrogen was identified as one of two residual donors in commercial ZnO samples; the presence of two donors in this material is consistent with Hall measurements. Hydrogen was found to correspond to the donor labeled $D1$, which has an ionization energy of 35 meV. The involvement of hydrogen in the structure of the $D1$ donor was confirmed by electron-nuclear double resonance (ENDOR). The $D1$ shallow donor produces a line with g values $g_{||} = 1.9569$ and $g_{\perp} = 1.9552$, consistent with the shallow-donor signals discussed in section 4.5.1. It is an interesting question whether these experiments observe the interstitial and/or the substitutional species. Because of the way the muons are introduced in the ZnO material, muon spin rotation probably observes the equivalent of isolated interstitial hydrogen.

Infrared (IR) spectroscopy has identified two lines corresponding to interstitial hydrogen in hydrogenated ZnO single crystals: one at 3326 cm^{-1} associated with the stretching H–O mode with H at the antibonding site [50, 53], shown in figure 8(*b*); and another at 3611 cm^{-1} associated with H at the bond-center site [203]. Shi *et al* recently showed that samples that were grown by a pressurized melt-growth technique showed a strong 3326 cm^{-1} line and a weak 3611 cm^{-1} line, and vice versa for samples grown by the seeded vapor transport technique [54]. The origin of these differences has been discussed by McCluskey and Jokela [55]. On the other hand, the identification of substitutional hydrogen by IR spectroscopy is much more difficult. The H_O vibrational modes occur in the range 750–1000 cm^{-1} [26]; in this spectral region ZnO samples tend to be opaque due to IR absorption by free carriers.

The experiments of Shi *et al* [54] indicate that hydrogen is stable in ZnO up to ∼500 °C, a result that is not consistent with the presence of the highly mobile interstitial hydrogen [51, 52]. However, the temperature dependence is consistent with the calculated migration barrier of 2.5 eV for substitutional hydrogen [26]. Moreover, H_O can also explain the observed dependence of the electrical conductivity on oxygen partial pressure, a dependence that, as we now know, has been erroneously attributed to the presence of oxygen vacancies [41].

6. Acceptor impurities

6.1. Lithium, sodium and potassium

Lithium may behave both as a donor and as an acceptor in ZnO [223]. The donor behavior arises when lithium occurs as an interstitial impurity; the acceptor behavior is exhibited when lithium substitutes on a Zn site. Kolb and Laudise [224] reported that incorporation of Li during growth can compensate the background n-type conductivity of hydrothermally grown ZnO, provided a post-growth heat treatment is applied. The heat treatment was considered necessary to achieve outdiffusion of zinc interstitials, but may also result in outdiffusion of Li interstitials as well as hydrogen donors. EPR studies of Li in ZnO were carried out by Kasai [170] and by Schirmer [225]. For the axial configuration, the measured g values were $g_{||} = 2.0028$ and $g_{\perp} = 2.0253$ [225]. Lithium has been used to compensate n-type doping; no p-type doping using Li has been reported.

Despite the size difference between Li and Na, it is expected that Na also shows an amphoteric behavior, acting as an acceptor when substituting for Zn and as a donor when occupying interstitial sites. The diffusion of Li and Na into ZnO single-crystal substrates, and subsequent characterization using mass spectroscopy and low-temperature photoluminescence measurements, have been reported by Meyer *et al* [226], and the results compared with Li- and Na-doped ZnO epitaxial films. In addition to behaving as deep acceptors, Meyer *et al* reported that Li and Na, incorporated either by diffusion or during thin-film growth, can also result in (relatively) shallow acceptors with binding energies around 300 meV, with the corresponding donor–acceptor pair recombinations showing a weak electron–phonon coupling [226].

Several groups have reported first-principles calculations for column-IA impurities (Li, Na and K) in ZnO [56, 58, 227, 228]. Park *et al* performed calculations for substitutional and interstitial Li, Na, and K in ZnO, and reported ionization energies of 0.09 eV, 0.17 eV and 0.32 eV for substitutional Li, Na and K, respectively [56]. They supported the notion that the p-type doping efficiency of column-IA elements is limited by the formation of compensating interstitials. Calculations by Wardle *et al* [58] indicated that the neutral Li_{Zn}–Li_i complex is relatively stable, with a dissociation energy of 1.5 eV. The activation of Li_{Zn} (i.e. dissociation of the Li_{Zn}–Li_i pair with subsequent outdiffusion Li_i^+) would occur at 300°. Wardle *et al* predicted an ionization/binding energy of 0.3 eV for Na_{Zn} and a dissociation energy of 0.9 eV for the Na_{Zn}–Na_i pair. Based on first-principles calculations Lee *et al* [227] proposed the

co-doping of Li and Na with H, in the form of hydrogen–acceptor complexes, as an effective way of enhancing the concentration of Li_{Zn} and Na_{Zn} acceptors. Activation of the hydrogen-passivated acceptors would be achieved by post-growth annealing.

Zeng *et al* [228] reported results for Li-doped p-type ZnO thin films prepared by dc reactive magnetron sputtering. They reported an activation energy for Li_{Zn} of 150 meV, extracted from free-to-neutral acceptor transitions. They also reported a deeper acceptor level at 250 meV, which was assigned to Li-related complexes. We note, however, that the reproducibility of these results by other groups remains an issue.

We note that first-principles calculations based on DFT within the local-density or generalized-gradient approximations present difficulties in describing the deep acceptor level that is experimentally observed for Li_{Zn}. Instead, DFT–LDA or GGA calculations have predicted Li_{Zn} to be a shallow acceptor, probably due to the inherent self-interaction error which favors a delocalized-hole picture [229]. Approaches that potentially overcome these difficulties, such as hybrid functionals, are yet to be applied to this problem.

6.2. Copper

Copper acts as a deep acceptor in ZnO. It can be used to reduce the carrier concentration in n-type ZnO by acting as a compensating center [230]. The electronic structure of the Cu impurity is complicated by the fact that a hole can be formed in the Cu 3d shell [231]. Copper has been mentioned as the source of green luminescence [196, 197]. It is likely that Cu is one of the causes, noting that not all samples that exhibit green luminescence are shown to contain Cu impurities.

6.3. Nitrogen

Among the acceptor impurities that substitute for oxygen in ZnO, nitrogen is thought to be the most suitable p-type dopant due to both atomic-size and electronic-structure considerations. The nitrogen atom is the closest in atomic size to oxygen and, therefore, it is expected to result in minimum strain in ZnO. The energy of the valence 2p states and the electronegativity of nitrogen are also the closest to those of the oxygen atom, particularly when compared with other column-V dopants [59]. Several groups have reported on the incorporation of N in ZnO, and many have claimed that N substitutes for O [60, 62, 69, 180, 232–237]; however, the reports on p-type conductivity in N-doped ZnO still remain controversial.

Nitrogen is a shallow acceptor in other II–VI semiconductors [238] and has been considered as a suitable p-type dopant for ZnO for some time [239]. Minegishi *et al* [60] reported p-type doping of ZnO films grown on sapphire (0 0 0 1) by chemical vapor deposition, using NH_3 as nitrogen source. They reported a carrier concentration of $1.5 \times 10^{16}\,cm^{-3}$ with an estimated ionization energy of 100 meV and Hall mobility of $12\,cm^2\,V^{-1}s^{-1}$. The authors also pointed out that hydrogen may play a role in the nitrogen incorporation, and that the appropriate growth/annealing conditions for obtaining p-type material are limited to such a narrow range that control turns out to be very difficult [60].

Thonke *et al* [232] performed photoluminescence spectroscopy measurements at low temperatures on ZnO bulk single crystals that were grown using the seeded chemical vapor transport (SCVT) technique by Eagle-Picher. They derived an acceptor binding energy of 195 meV based on an analysis of donor–acceptor pair transitions, and assigned it to N_O. Note that the ZnO single crystals in the experiments by Thonke *et al* were not intentionally doped, but were supposed to contain nitrogen. Carlos *et al* [180] have reported EPR signals for N_O in ZnO single crystals from Eagle-Picher, finding g values $g_{||} = 1.9953$ and $g_{\perp} = 1.9633$. These values seem to be in reasonable agreement with the observations by Garces *et al* [181] who reported $g_{||} = 1.9948$ and $g_{\perp} = 1.9632$. The large anisotropy in the hyperfine interaction was attributed by Carlos *et al* to a Jahn–Teller distortion along the c axis away from the substitutional site, that would make N a deep acceptor in ZnO [180].

Guo *et al* [233] reported the growth of LED structures with nitrogen-doped ZnO films using plasma-enhanced laser deposition. The films were grown on n-type ZnO single-crystal wafers, using N_2O as nitrogen source in the N-doped ZnO layer. Four-probe van der Pauw measurements showed a relatively high resistivity of 100–200 Ω cm, which led Guo *et al* to suggest that N introduces a relatively deep level in ZnO. The observed bluish emission observed in the electroluminescence measurements by Guo *et al* is consistent with sub-band-gap recombinations.

Look *et al* [62] investigated the electrical and optical properties of N-doped ZnO grown by molecular-beam epitaxy on Li-diffused semi-insulating ZnO bulk single crystals from Eagle-Picher. Nitrogen was incorporated by using a flux of N_2 gas added to the O_2 gas flow in the rf plasma source. Hall-effect and conductivity measurements resulted in a hole concentration of $9 \times 10^{16}\,cm^{-3}$, mobility of $2\,cm^2\,V^{-1}\,s^{-1}$, and resistivity of 40 Ω cm. Photoluminescence measurements at 2 K revealed a peak at 3.315 eV, which was attributed to a bound-exciton to neutral–acceptor recombination. Based on the band gap of 3.437 eV for ZnO at 2 K and a donor binding energy of 60 meV, and a LO-phonon replica at 3.238 eV supposedly related to the 3.315 eV peak, Look *et al* estimated acceptor ionization energy to be 0.17–0.20 eV. They also noted that the high ionization energy and relatively low hole concentration are consistent with secondary-ion mass spectroscopy measurements which showed N concentrations of about $10^{19}\,cm^{-3}$ in the N-doped films, approximately two orders of magnitude higher than the ionized acceptor concentration [62].

Meyer *et al* [234] performed photoluminescence measurements in nitrogen-implanted ZnO bulk crystals from Eagle-Picher and from CRYSTEC, as well as N-doped ZnO epitaxial films grown by vapor phase deposition (CVD) using NH_3 as the nitrogen source. They reported the appearance of a donor–acceptor recombination at 3.235 eV after implantation and activation (annealing at 950 °C), instead of a peak at 3.315 eV as reported by Look *et al* [62]. Based on the 3.235 eV peak, Meyer *et al* derived an acceptor binding energy of

165 meV for N in ZnO. They also reported the appearance of a second donor–acceptor recombination at 3.22 eV in the N-doped ZnO films, which was attributed to an unknown residual acceptor [234]. More recently, Lautenschlaeger *et al* [235] investigated the nitrogen incorporation in homoepitaxial ZnO layers grown by chemical vapor deposition, using NH_3 as nitrogen precursor. They explored the growth on both Zn-polar and O-polar faces, and found the nitrogen incorporation to be more favorable on the Zn-polar face. Based on Raman and photoluminescence spectroscopy measurements, for tracing the N incorporation, they concluded that relatively low growth temperatures (∼400 °C) and Zn-polar single-crystal substrates are essential for N incorporation in homoepitaxy using chemical vapor deposition techniques [235].

Tsukazaki *et al* [69] reported the realization of a ZnO LED by laser MBE on insulating $ScAlMgO_4$. The p-type N-doped layer was grown by a repeated temperature modulation technique in which 15 nm high-N-concentration layers were grown at low temperatures (∼450 °C), then annealed at higher temperatures, followed by 1 nm low-N-concentration layers at high temperatures (∼1050 °C). The whole process was intended to lead to an overall high N incorporation in the N-doped layer. Following this procedure, Tsukazaki *et al* reported blue light emission from a p–n ZnO homojunction using N as the acceptor in the p-type layer. They reported hole mobility and acceptor activation energy of 5–8 $cm^2\ V^{-1}s^{-1}$ and 100 meV, respectively. However, it is important to note that the band gap of ZnO is 3.4 eV and, as such, one should expect UV light emission from a ZnO p–n junction. As in the case of previous reports on N-doped ZnO, the work of Tsukazaki *et al* [69] was not followed up with further developments, leading to doubts on the reliability and reproducibility of the p-type doping.

Recently, Fang *et al* [236] performed deep-level transient spectroscopy measurements on a *n*-ZnO/*i*-ZnO/ZnO : N structure grown on a p-type Si substrate using metalorganic chemical vapor deposition for the growth of the ZnO : N layer and sputtering deposition for the growth of the *i*-ZnO and n^+-ZnO : Al layers. Their results led them to propose the existence of an electron trap at 0.29 eV and a hole trap at 0.9 eV in the ZnO : N layer. The nature of these traps, i.e. whether the trap at 0.29 eV is related to native defects or the trap at 0.9 eV is related to N, is still unknown.

The compensation of nitrogen substitutional acceptors by intrinsic defects as well as by nitrogen incorporated in different configurations has been theoretically investigated by Lee *et al* [19]. Based on calculations using DFT within the LDA, they concluded that a low N doping level would result from N incorporation using a normal N_2 source, with the N acceptors being compensated mainly by oxygen vacancies. At high doping levels, N acceptors would be compensated via the formation of defect complexes with Zn antisites. They also added that even if a N_2 plasma is used to increase the nitrogen incorporation (due to an increased N chemical potential), N acceptors would still be compensated by N_2 molecules at oxygen sites [$(N_2)_O$] and N_O–$(N_2)_O$ complexes, which would explain the difficulty in achieving low-resistivity p-type ZnO [19]. On the other hand, recent first-principles calculations by Fons *et al* [237] indicate that nitrogen does not predominantly incorporate in the form of $(N_2)_O$ when a plasma source is used. Instead, a comparison between N K-edge x-ray absorption spectra and results from first-principles calculations indicates that nitrogen incorporates substitutionally on O sites, where it is expected to act as an acceptor. Moreover, annealing at relatively high temperatures (800 °C) may lead to the formation of N_2 bubbles [237], suggesting that the incorporation of high concentrations of 10^{20} cm^{-3} N_O in ZnO can be achieved only in a metastable manner.

Despite the results of Lee *et al* [19], it is important to note that the mechanism of compensation of N_O acceptors in ZnO is still under debate. For instance, the $\varepsilon(0/-)$ transition level (or ionization/binding energy) of N_O at 0.44 eV above the valence-band maximum as predicted by Lee *et al* is too high for a shallow acceptor. More importantly, this carries a large uncertainty, mainly due to the band gap underestimation by the standard DFT-LDA approach [19]. That is, even the shallow/deep nature of the N_O acceptor is not well established, and more sophisticated methods will be necessary to elucidate the electronic structure of N in ZnO.

6.4. Phosphorus, arsenic and antimony

Phosphorus and arsenic have been found to introduce *deep* acceptor states in ZnSe [240], so the prospects for these impurities to act as shallow acceptors in ZnO are not good. Antimony is expected to result in even deeper states. Aoki *et al* [63] have reported ZnO diodes in which a p-type layer was created by excimer laser doping from a Zn_3P_2 compound. No details about the acceptors were reported. Ryu *et al* have investigated arsenic as an acceptor in ZnO grown on GaAs(0 0 1) [65]. Hall measurements indicated p-type conductivity, but it was pointed out that large uncertainties resulted due to contributions from 'interference layers' between the ZnO film and the GaAs substrate. From optical measurements an ionization energy of 100 meV was deduced. Chu *et al* reported results of Sb-doped ZnO grown on Si, with a relatively high hole concentration of 2×10^{18} cm^{-3} [67]. The same group recently reported results for a p–n homojunction using Sb as p-type dopant, with a turn-on voltage of 6 eV [68]. Based on first-principles calculations, Limpijumnong *et al* proposed that these large anions (As, Sb) occupy Zn lattice sites and form complexes with two nearby Zn vacancies [70]. A similar model has also been proposed for P in ZnO by Lee *et al* [241]. These calculations indicate that these complexes are shallow donors with ionization energies less than 100 meV. However, the formation of these complexes is quite unlikely from an energetic point of view, as discussed in section 4.6. Entropic considerations also argue against the formation of complexes comprised of three constituents.

6.5. Co-doping

Co-doping has been suggested to be an effective method of achieving p-type conductivity in ZnO. The term co-doping means that, along with the acceptors that are incorporated to produce holes, donors are also incorporated during the growth. At first sight, this would lead merely to compensation. In fact,

compensation during the growth is actually quite desirable, since it shifts the Fermi level away from the VBM toward the middle of the gap. This results in a lowering of the formation energy of acceptors (and hence in an increase in the acceptor solubility), as well as an increase in the formation energy of compensating donor-type native defects (such as V_O). However, the compensation by the intentionally introduced donor will persist after growth, and the material will not exhibit p-type conductivity.

One potential strategy for overcoming this limitation is to remove the donor after growth. This is only possible with donor impurities that are not strongly bound and that exhibit a sufficiently high diffusivity, so that they can be removed from the vicinity of the acceptors during an anneal at modest temperatures (to avoid formation of other compensating defects). It has been proposed that hydrogen may be a candidate for such a donor [16, 17]. Indeed, hydrogen plays such a role in p-type doping of GaN, where it enhances the solubility of Mg acceptors and suppresses compensation by nitrogen vacancies [242]. Whether this type of dopant engineering also works in the case of ZnO will depend on the binding and dissociation energies of acceptor–hydrogen complexes, and on the barriers that need to be overcome to remove H from the vicinity of the acceptors during a post-growth anneal. The potential beneficial aspects of simultaneous incorporation of hydrogen along with nitrogen acceptors were noted by Minegishi *et al* [60].

Another type of co-doping has been proposed by Yamamoto and Katayama-Yoshida [243]. This proposal is based on the incorporation of complexes consisting of two acceptors and one donor. The authors claimed that such complexes would result in higher hole concentrations due to an enhancement of the solubility and a lowering of the ionization energy. These conclusions were qualitatively based on first-principles calculations, but formation energies and ionization energies were not calculated. Explicit calculations for similar situations, i.e. for acceptor doping of GaN [244] and of CdTe [245] have in fact indicated that co-doping with such complexes is unlikely to produce favorable results.

Experimentally, Joseph *et al* [61] have reported p-type ZnO through codoping using N (in the form of N_2O) with Ga as a codopant. They reported suspiciously high hole concentrations of $4 \times 10^{19}\,\mathrm{cm}^{-3}$ and low room temperature resistivity of $2\,\Omega\,\mathrm{cm}$. Yan *et al* [246] have proposed an alternative interpretation of the results, in terms of the chemical potentials of the gases used during growth. More recently, Yan *et al* [247] suggested that co-doping would be effective in the case of ZnO only for very high Ga and N concentrations so that an impurity band (caused by neutral Ga + N complexes) would form above the ZnO valence band, leading to a significant reduction of the ionization energy of the excess N impurities. We note, however, that the experimental results for co-doping in ZnO have not been reproduced.

7. MgZnO and CdZnO alloys for band-gap engineering

In addition to controllable n-type and p-type doping, the success of ZnO as a semiconductor also depends on the possibilities of band-gap engineering. Most semiconductor device designs rely on heterostructures for providing carrier and/or optical confinement; examples include LEDs and lasers based on quantum wells, and high electron mobility transistors (HEMTs). These heterostructures are composed of layers of different materials/compositions, and the most relevant parameters in this case are the band gaps of each layer and the valence- and conduction-band offsets between the individual layers. Band-gap engineering in ZnO can be achieved by alloying with MgO and/or CdO, in analogy to GaN which can be alloyed with AlN and/or InN. One of the differences between the nitride and the oxide families is that AlN and InN assume the same wurtzite crystal structure of GaN, whereas MgO and CdO assume the rock-salt structure, not the same as the ZnO wurtzite structure. This is indeed a problem for ZnO-based alloys with high content of MgO or CdO, in which case phase separation is expected to occur. For moderate contents of MgO and CdO, however, MgZnO and CdZnO alloys assume the wurtzite crystal structure of the parent compound, still providing a wide range of band gaps (2.3–4.0 eV) and band offsets for practical device applications [113–118].

7.1. Synthesis and characterization of MgZnO and CdZnO alloys

MgZnO films have been grown on a variety of substrates, including sapphire (0 0 0 1), GaN/sapphire (0 0 0 1), $ScAlMgO_4$ and ZnO, using different growth techniques, such as pulsed laser deposition, metal-organic vapor-phase epitaxy and molecular-beam epitaxy [113–118]. It has been found that MgZnO films assume the wurtzite structure for Mg concentrations up to ∼33%; a phase transition to the cubic structure has been observed for higher concentrations [113]. A band-gap increase up to 3.99 eV for 33% Mg has been reported [113]. X-ray diffraction measurements indicate that the lattice parameter a increases and c decreases with increasing Mg concentration [113]. An increase in the exciton binding energy has been observed for MgZnO alloys [117].

CdZnO alloys with Cd concentrations of 7% and band gap of 3.0 eV were grown on sapphire (0 0 0 1) and $ScAlMgO_4$ substrates using pulsed-laser deposition [116]. While adding Mg to ZnO leads to an increase in the band gap, adding Cd leads to a decrease in the band gap, similar to the case of addition of Al and In to GaN, respectively. In the case of CdZnO, both lattice parameters a and c increase with the Cd content, consistent with the larger atomic size of Cd compared with Zn. CdZnO films were also grown by metal-organic vapor-phase epitaxy on sapphire (0 0 0 1) [118]. A large redshift of 300 meV in the band gap for a lattice mismatch of only 0.5% with respect to ZnO has been observed. These samples, however, were shown to exhibit large inhomogeneities, containing laterally distinguishable regions with different Cd concentrations [118].

Wassner *et al* [248] reported the growth of wurtzite MgZnO thin films with Mg concentrations of up to 37% on sapphire (0 0 0 1) by plasma-assisted molecular-beam epitaxy, and using MgO/MgZnO as a buffer layer. They found the lattice parameter a to be independent of Mg concentration, whereas c decreased from 5.20 Å for ZnO to 5.17 Å for Mg

concentration of 37%, indicating a pseudomorphic growth. They observed a peak position of band-edge luminescence at 4.11 eV for 37% of Mg. Low-temperature optical transmission measurements showed strong excitonic features for all Mg compositions, with alloy broadening effects observed for higher Mg content. They observed an increase in the Stokes shift of 125 meV and in the Urbach energy of 54 meV, as extracted from the exponential tail in the sub-band-gap region of the absorption spectrum for Mg concentration of 37%. This was interpreted as resulting from exciton localization due to fluctuations in the alloy concentration [248].

Lautenschlaeger *et al* [249] reported the growth of MgZnO epilayers on ZnO single-crystal substrates using chemical vapor deposition at relatively low temperatures of ∼650 °C, without using organometallic precursors. Instead, they used metallic precursors (Zn, Mg) along with NO_2 as an oxygen precursor. Epitaxial films with fixed Mg concentrations were found to require an individual adjustment of the Mg reservoir temperature, allowing Lautenschlaeger *et al* to explore the possibility of achieving a Mg gradient within one sample in a combinatorial approach using chemical vapor deposition [249].

The effect of Mg substitution on the crystal structure of wurtzite ZnO was investigated with synchrotron x-ray diffraction by Kim *et al* [250–252]. An increase in Mg concentration in polycrystalline wurtzite $Zn_{1-x}Mg_xO$ (x <0.15) samples (prepared by an oxalate-based coprecipitation method) resulted in a pronounced c-axis compression of the wurtzite lattice, and a decrease in the off-center cation displacement within each tetrahedral ZnO_4 unit. Significant changes in the ionic polarization were observed (from $-5.6\,\mu C\,cm^{-2}$ for $x = 0$ to $-4.8\,\mu C\,cm^{-2}$ for $x = 0.15$), despite only subtle increments in the cell volume (∼0.03%) and the planar areal dimension (∼0.1%). In addition to synchrotron x-ray diffraction, Raman spectroscopy, NMR and neutron scattering were also applied [251, 252]; the results suggest that Mg substitution will increase the spontaneous polarization of ZnO.

7.2. Heterostructures

Using MgZnO as barrier layers, Ohtomo *et al* observed excitonic stimulated emission from ZnO/MgZnO superlattices up to 373 K; the emission energy varied from 3.2 to 3.4 eV, depending on the well thickness and/or the Mg content in the barrier layers [114].

Quantum effects of a high-mobility two-dimensional electron gas have been observed in polar ZnO/MgZnO heterostructures grown by laser molecular-beam epitaxy on $ScAlMgO_4$ substrates. The observed sheet electron density varied from 0.7×10^{12} to $3.7 \times 10^{12}\,cm^{-2}$ depending on the Mg content in the barriers and the growth polarity [253]. A two-dimensional electron gas in ZnO/MgZnO grown on sapphire (0 0 0 1) by molecular-beam epitaxy, with sheet carrier density of $1.1 \times 10^{13}\,cm^{-2}$ has also been observed [254]; the polarization of the wurtzite structure along the c-direction favorably affects the formation of the 2DEG in ZnO/MgZnO, in analogy to the 2DEG observed in AlGaN/GaN interfaces [254].

Growth of MgZnO films by plasma-assisted molecular-beam epitaxy on Zn-polar ZnO substrates was recently reported [255]. Time-resolved photoluminescence measurements were used to characterize the structural quality of the samples; photoluminescence lifetimes on the order of nanoseconds were observed, which are indicative of a relatively low concentration of non-radiative defects [255]. The growth of non-polar MgZnO/ZnO heterostructures along the a-direction [11–20] on sapphire by molecular-beam epitaxy has also been reported [256]. The non-polar quantum wells exhibit quantum confinement without any indication of quantum Stark effects, in contrast to the heterostructures grown along the polar c-direction [256].

In the design of devices based on heterostructures, the key parameters are the valence-band and conduction-band offsets between the individual layers. These layers are often strained, requiring the knowledge of the effects of strain on the band structure; these effects are described by *deformation potentials*. These quantities are difficult to extract from experiments alone, so that first-principles calculations are very helpful. Once these quantities are obtained for binary compounds, band offsets and deformation potentials for alloys can be obtained from the respective values for the parent compounds by interpolation [257]. An additional difficulty in obtaining these quantities occurs in the case of MgZnO and CdZnO alloys: MgO and CdO do not exist in the wurtzite structure, so that the interpolation scheme based on values for the parent compounds does not apply. To apply interpolation, values for the wurtzite phase of MgO and CdO are required, and such values cannot be obtained from experiment. In this case, first-principles calculations for the hypothetical wurtzite phases of MgO and CdO are particularly helpful.

7.3. Deformation potentials and band alignments in MgZnO and CdZnO alloys

Knowing how strain affects the band structure of MgZnO and CdZnO alloys and the band alignment or offset between these materials and the parent compound ZnO is essential for the design of optoelectronic devices that rely on band-gap engineering to provide carrier and optical confinement. Band offsets at strained-layer interfaces are obtained based on offsets between the unstrained layers by adding the effects of strain through the use of deformation potentials. These potentials describe the splitting of bands under uniaxial or biaxial strain, and also the overall shift of individual band edges under hydrostatic strain. The latter are expressed by the 'absolute' deformation potentials a_v for the valence band and a_c for the conduction band. While shifts of the conduction band with respect to the valence band can be straightforwardly extracted from experimental measurements, the absolute deformation potentials are much more difficult to obtain. Computationally, their derivation is also complicated due to the lack of an absolute reference for the potential in a homogeneously deformed crystal.

Recently, first-principles calculations of a_v and a_c for wurtzite ZnO, MgO and CdO, and band offsets between the MgO/ZnO and ZnO/CdO interfaces, have been reported [258]. Instead of explicitly calculating a_v and a_c and band offsets for specific alloy compositions, it is a common practice to derive

Table 2. Calculated absolute deformation potentials (in eV) of wurtzite MgO, ZnO, CdO. a_v is the deformation potential for the VBM and a_c for the CBM. $a_g = a_c - a_v$ denotes the (relative) band-gap deformation potential.

	MgO	ZnO	CdO
a_v	2.0	−0.2	0.0
a_c	−4.3	−3.1	−0.4
a_g	−6.3	−2.9	−0.4

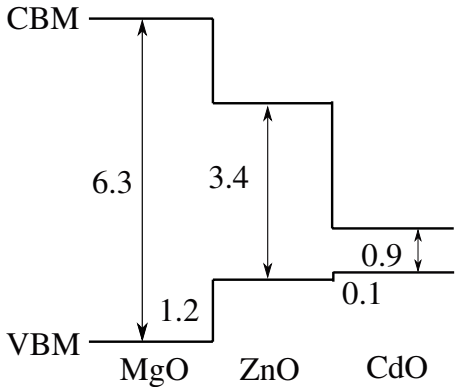

Figure 10. Calculated band offsets at wurtzite MgO/ZnO and ZnO/CdO interfaces. All values are in eV (From [258]. With permission.)

values for alloys by interpolating the values for the parent compounds. The calculated values for a_v and a_c in ZnO, MgO and CdO are listed in table 2. Compared with conventional semiconductors or III-nitrides [259, 260], the oxides generally have smaller band-gap deformation potentials. The absolute deformation potential a_v is very small for ZnO and CdO due to the proximity of the semicore d bands to the VBM [258].

The calculated band alignment at MgO/ZnO and ZnO/CdO interfaces is shown in figure 10. The large valence-band offset at the MgO/ZnO is due to the absence of d bands in MgO. Due to the large ratio of the conduction-band to the valence-band offsets, it is expected that most of the change in the band gap of CdZnO alloys will be accommodated by the variation in the conduction-band edge. Coupled with the very small valence-band deformation potentials of ZnO and CdO, we expect very weak confinement of holes in ZnO/CdZnO/ZnO heterostructures, even for relatively high Cd concentrations.

Because the values for MgO and CdO reported in table 2 and figure 10 are for the wurtzite phase, values for MgZnO and CdZnO alloys can be obtained by interpolation for the experimentally relevant range of alloy compositions over which the alloys exhibit the ZnO-matching wurtzite structure.

8. Summary and outlook

Despite the difficulties in achieving p-type conductivity, ZnO remains a promising material for electronic and optoelectronic device applications. The difference between the current surge to turn ZnO into a functional semiconductor and past attempts is that better thin-film growth techniques are available, ZnO substrates have higher quality and experience has been gained from research on III-nitrides, which posed similar difficulties with respect to controlling the conductivity. On the computational side, we have access to faster computers and improved first-principles methodologies, so that the electronic, structural and optical properties of native defects, impurities, surfaces and interfaces can be predicted with higher accuracy. Computations have made major contributions to understanding the electronic structure of oxygen vacancies and the shallow-donor nature of hydrogen, and to screening possible shallow acceptors in ZnO.

A key conclusion from the recent studies is that native point defects cannot explain the often-observed n-type conductivity, but the latter is likely to be caused by the incorporation of impurities during growth or annealing. It may still take some time for the community to adjust to this point of view, but the precedent of GaN may be informative. Until the early 1990s, the overwhelming consensus in the field was that GaN exhibited unintentional n-type conductivity due to nitrogen vacancies, and that p-type GaN would be extremely difficult or even impossible to achieve. Computational studies [261, 262] then showed that nitrogen vacancies are high-energy defects that are highly unlikely to occur in n-type GaN. Instead, n-type conductivity was attributed to unintentional impurities such as oxygen. This view proved to be correct and within a few years the perception of the role of point defects completely changed, and the focus was shifted to the importance of controlling impurity incorporation. ZnO will likely follow this historical precedent.

ZnO with controlled n-type conductivity has many important applications, such as transparent contacts and high electron mobility transistors (HEMTs). If reliable and reproducible p-type doping can be achieved, it would hugely boost the applications of ZnO, for instance in LEDs and lasers. An important factor in these developments is that the quality and availability of ZnO substrates have dramatically improved in recent years, and that high-quality epitaxial layers are now being controllably produced. Still, a number of areas still need more investigations; in particular, we note the possible presence of an electron accumulation layer on the surface, and its effects on measurements of conductivity of the underlying layers. A pragmatic synergism between theory and experiment can greatly aid in achieving the desired objectives.

Acknowledgments

Thanks are due to L Halliburton, M D McCluskey, J Neugebauer, T Ive, J Speck and G D Watkins for valuable discussions. This work was supported by the NSF MRSEC Program under award No DMR05-20415 and by the UCSB Solid State Lighting and Energy Center.

References

[1] Look D C 2001 *Mater. Sci. Eng.* B **80** 383
[2] Özgür Ü, Alivov Y I, Liu C, Teke A, Reshchikov M A, Dogan S, Avrutin V, Cho S-J and Morkoç H 2005 *J. Appl. Phys.* **98** 041301
[3] Ogale S B 2005 *Thin Films and Heterostructures for Oxide Electronics* (New York: Springer)

[4] Nickel N H and Terukov E (ed) 2005 *Zinc Oxide—A Material for Micro- and Optoelectronic Applications* (Netherlands: Springer)
[5] Jagadish C and Pearton S J (ed) 2006 *Zinc Oxide Bulk, Thin Films, and Nanostructures* (New York: Elsevier)
[6] Thomas D G 1960 *J. Phys. Chem. Solids* **15** 86
[7] Mang A, Reimann K and Rübenacke St 1995 *Solid State Commun.* **94** 251
[8] Reynolds D C, Look D C, Jogai B, Litton C W, Cantwell G and Harsch W C 1999 *Phys. Rev.* B **60** 2340
[9] Chen Y, Bagnall D M, Koh H-J, Park K-T, Hiraga K, Zhu Z-Q and Yao T 1998 *J. Appl. Phys.* **84** 3912
[10] Srikant V and Clarke D R 1998 *J. Appl. Phys.* **83** 5447
[11] Reynolds D C, Look D C and Jogai B 1996 *Solid State Commun.* **99** 873
[12] Bagnall D M, Chen Y F, Zhu Z, Yao T, Koyama S, Shen M Y and Goto T 1997 *Appl. Phys. Lett.* **70** 2230
[13] Brown M E (ed) 1957 *ZnO—Rediscovered* (New York: The New Jersey Zinc Company)
[14] Van de Walle C G 2000 *Phys. Rev. Lett.* **85** 1012
[15] Kohan A F, Ceder G, Morgan D and Van de Walle C G 2000 *Phys. Rev.* B **61** 15019
[16] Van de Walle C G 2001 *Physica* B **308–310** 899
[17] Zhang S B, Wei S-H and Zunger A 2001 *Phys. Rev.* B **63** 075205
[18] Oba F, Nishitani S R, Isotani S, Adachi H and Tanaka I 2001 *J. Appl. Phys.* **90** 824
[19] Lee E-C, Kim Y-S, Jin Y-G and Chang K J 2001 *Phys. Rev.* B **64** 085120
[20] Janotti A and Van de Walle C G 2005 *Appl. Phys. Lett.* **87** 122102
[21] Erhart P, Klein A and Albe K 2005 *Phys. Rev.* B **72** 085213
[22] Janotti A and Van de Walle C G 2006 *J. Cryst. Growth* **287** 58
[23] Erhart P and Albe K 2006 *Phys. Rev.* B **70** 115207
[24] Lany S and Zunger A 2005 *Phys. Rev.* B **72** 035215
Lany S and Zunger A 2007 *Phys. Rev. Lett.* **98** 045501
[25] Patterson C H 2006 *Phys. Rev. Lett.* **74** 144432
[26] Janotti A and Van de Walle C G 2007 *Phys. Rev.* B **6** 44
[27] Janotti A and Van de Walle C G 2007 *Phys. Rev.* B **75** 165202
[28] Paudel T R and Lambrecht W R L 2008 *Phys. Rev.* B **77** 205202
[29] Oba F, Togo A, Tanaka I, Paier J and Kresse G 2008 *Phys. Rev.* B **77** 245202
[30] Nakamura S and Chichibu S F (ed) 2000 *Nitride Semiconductor Blue Lasers and Light Emitting Diodes* (Boca Raton, FL: CRC Press)
[31] Lannoo M and Bourgoin J 1981 *Point Defects in Semiconductors I: Theoretical Aspects* (Berlin: Springer)
Lannoo M and Bourgoin J 1983 *Point Defects in Semiconductors II: Experimental Aspects* (Berlin: Springer)
[32] Pantelides S T (ed) 1992 *Deep Centers in Semiconductors: A State-of-the-Art Approach* 2nd edn (Yverdon: Gordon and Breach)
[33] Stavola M (ed) 1999 *Identification of Defects in Semiconductors*, *Semiconductors and Semimetals* vol 51B (San Diego: Academic)
[34] Harrison S E 1954 *Phys. Rev.* **93** 52
[35] Hutson A R 1957 *Phys. Rev.* **108** 222
[36] Thomas D G 1957 *J. Phys. Chem. Solids (USA)* **3** 229
[37] Mohanty G P and Azároff L V 1961 *J. Chem. Phys.* **35** 1268
[38] Hausmann A 1970 *Z. Phys.* **237** 86
[39] Hausmann A and Tuerle W 1973 *Z. Phys.* **259** 189
[40] Hoffmann K and Hahn D 1974 *Phys. Status Solidi* a **24** 637
[41] Kröger F A 1974 *The Chemistry of Imperfect Crystals* (Amsterdam: North-Holland)
[42] Utsch B and Hausmann A 1975 *Z. Phys.* B **21** 27
[43] Hausmann A and Utsch B 1975 *Z. Phys.* B **21** 217
[44] Hagemark K I 1976 *J. Solid State Chem.* **16** 293
[45] Neumann G 1981 *Current Topics in Materials Science* vol 7 ed E Kaldis (Amsterdam: North Holland) p 152
[46] Vlasenko L S and Watkins G D 2005 *Phys. Rev.* B **72** 035203
[47] Vlasenko L S and Watkins G D 2005 *Phys. Rev.* B **71** 125210
[48] Cox S F J *et al* 2001 *Phys. Rev. Lett.* **86** 2601
[49] 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
[50] McCluskey M D, Jokela S J, Zhuravlev K K, Simpson P J and Lynn K G 2002 *Appl. Phys. Lett.* **81** 3807
[51] Thomas D G and Lander J J 1956 *J. Chem. Phys.* **25** 1136
[52] Wardle M G, Goss J P and Briddon P R 2006 *Phys. Rev. Lett.* **96** 205504
[53] Jokela S J and McCluskey M D 2005 *Phys. Rev.* B **72** 113201
[54] Shi G A, Stavola M, Pearton S J, Thieme M, Lavrov E V and Weber J 2005 *Phys. Rev.* B **72** 195211
[55] McCluskey M D and Jokela S J 2007 *Physica* B **401–402** 355
[56] Park C H, Zhang S B and Wei S-H 2002 *Phys. Rev.* B **66** 073202
[57] Lee E C and Chang K J 2004 *Phys. Rev.* B **70** 115210
[58] Wardle M G, Goss J P and Briddon P R 2005 *Phys. Rev.* B **71** 155205
[59] Harrison W A 1999 *Elementary Electronic Structure* (Singapore: World Scientific)
[60] Minegishi K, Koiwai Y, Kikuchi Y, Yano K, Kasuga M and Shimizu A 1997 *Japan. J. Appl. Phys. Part* II **36** L1453
[61] Joseph M, Tabata H and Kawai T 1999 *Japan. J. Appl. Phys. Part* II **38** L1205
[62] Look D C, Reynolds D C, Litton C W, Jones R L, Eason D B and Cantwell G 2002 *Appl. Phys. Lett.* **81** 1830
[63] Aoki T, Hatanaka Y and Look D C 2000 *Appl. Phys. Lett.* **76** 3257
[64] Kim K K, Kim H S, Hwang D K, Lim J H and Park S J 2003 *Appl. Phys. Lett.* **83** 63
[65] Ryu Y R, Zhu S, Look D C, Wrobel J M, Yeong H M and White H W 2000 *J. Cryst. Growth* **216** 330
[66] Ryu Y R, Lee T S and White H W 2003 *Appl. Phys. Lett.* **83** 87
[67] Xiu F X, Yang Z, Mandalapu L J, Zhao D T and Liu J L 2005 *Appl. Phys. Lett.* **87** 152101
[68] Chu S, Lim J H, Mandalapu L J, Yang Z, Li L and Liu J L 2005 *Appl. Phys. Lett.* **92** 152103
[69] Tsukazaki A *et al* 2005 *Nature Mater.* **4** 42
[70] Limpijumnong S, Zhang S B, Wei S H and Park C H 2004 *Phys. Rev. Lett.* **92** 155504
[71] Schmidt O, Kiesel P, Van de Walle C G, Johnson N M, Nause J and Döhler G H 2005 *Japan. J. Appl. Phys. Part 1* **44** 7271
[72] Schmidt O, Geis A, Kiesel P, Van de Walle C G, Johnson N M, Bakin A, Waag A and Döhler G H 2006 *Superlatt. Microstruct.* **39** 8
[73] Look D C 2007 *Surf. Sci.* **601** 5315
[74] Ohgaki T, Ohashi N, Sugimura S, Ryoken H, Sakaguchi I, Adachi Y and Haneda H 2008 *J. Mater. Res.* **23** 2293
[75] Bierwagen O, Ive T, Van de Walle C G and Speck J S 2008 *Appl. Phys. Lett.* **93** 242108
[76] Look D C, Reynolds D C, Sizelove J R, Jones R L, Litton C W, Cantwell G and Harsch W C 1998 *Solid State Commun.* **105** 399
[77] Maeda K, Sato M, Niikura I and Fukuda T 2005 *Semicond. Sci. Technol.* **20** S49
[78] Triboulet R, Munoz-Sanjos V, Tena-Zaera R, Martinez-Tomas M C and Hassani S 2005 *Zinc Oxide—A Material for Micro- and Optoelectronic Applications (Nato Science Series)* ed N H Nickel and E Terukov (The Netherlands: Springer)
[79] Ohtomo A and Tsukasaki A 2005 *Semicond. Sci. Technol.* **20** S1

[80] Heinze S, Krtschil A, Blasing J, Hempel T, Veit P, Dadgar A, Christen J and Krost A 2007 *J. Cryst. Growth* **308** 170
[81] Dadgar A *et al* 2004 *J. Cryst. Growth* **272** 800
[82] Ive T, Ben-Yaacov T, Van de Walle C G, Mishra U K, DenBaars S P and Speck J S 2008 *J. Cryst. Growth* **310** 3407
[83] Ive T, Ben-Yaacov T, Murai A, Asamizu H, Van de Walle C G, Mishra U, DenBaars S P and Speck J S 2008 *Physica Status Solidi* c **5** 3091
[84] Madelung O (ed) 1996 *Semiconductors—Basic Data 2nd Revised Edn* (Berlin: Springer)
[85] Cao H, Zhao Y G, Ho S T, Seelig E W, Wang Q H and Chang R P H 1999 *Phys. Rev. Lett.* **82** 2278
[86] Huang M H, Mao S, Feick H, Yan H, Wu Y, Kind H, Weber E, Russo R and Yang P 2001 *Science* **292** 1897
[87] Ryu Y R, Kim W J and White H W 2000 *J. Cryst. Growth* **219** 419
[88] Mandalapu L J, Yang Z, Xiu F X, Zhao D T and Liu J L 2006 *Appl. Phys. Lett.* **88** 092103
[89] Chu S, Olmedo M, Yang Z, Kong J and Liu J L 2008 *Appl. Phys. Lett.* **93** 181106
[90] Yu P Y and Cardona M 2005 *Fundamentals of Semiconductors* 3rd edn (Berlin: Springer)
[91] Kamalasanan M N and Chandra S 1996 *Thin Solid Films* **288** 112
[92] Paraguay F D, Estrada W L, Acosta D R N, Andrade E and Miki-Yoshida M 1999 *Thin Solid Films* **350** 192
[93] Funakubo H, Mizutani N, Yonetsu M, Saiki A and Shinozaki K 1999 *J. Electroceram.* **4:S1** 25
[94] Sakurai K, Kanehiro M, Nakahara K, Tanabe T, Fujita S and Fujita S 2000 *J. Cryst. Growth* **209** 522
[95] Yamamoto T, Shiosaki T and Kawabata A 1980 *J. Appl. Phys.* **51** 3113
[96] Molarius J, Kaitila J, Pensala T and Ylimlammi M 2003 *J. Mater. Sci.: Mater. Electron.* **14** 431
[97] Ondo-Ndong R, Ferblantier G, Pascal-Delannoy F, Boyer A and Foucaran A 2003 *Microelectron. J.* **34** 1087
[98] Gardeniers J G E, Rittersma Z M and Burger G J 1998 *J. Appl. Phys.* **83** 7844
[99] Shionoya S and Yen W H (ed) 1997 *Phosphor Handbook By Phosphor Research Society* (Boca Raton, FL: CRC Press)
[100] Nanto H, Sokooshi H and Usuda T 1991 *Solid-State Sensors and Actuators* **24–27** 596
[101] Eda K 1989 *Electr. Insul. Mag. IEEE* **5** 28
[102] Larciprete M C *et al* 2006 *Appl. Phys. B—Lasers Opt.* **82** 431
[103] Florescu D I, Mourokh L G, Pollak F H, Look D C, Cantwell G and Li X 2002 *J. Appl. Phys.* **91** 890
[104] Özgür Ü, Gu X, Chevtchenko S, Spradlin J, Cho S-J, Morkoç H, Pollak F H, Everitt H O, Nemeth B and Nause J E 2006 *J. Electr. Mater.* **35** 550
[105] Suscavag M *et al* 1999 *MRS Internet J. Nitride Semicond. Res.* **4S1** G3.40
[106] Ohshima E, Ogino H, Niikura I, Maeda K, Sato M, Ito M and Fukuda T 2004 *J. Cryst. Growth* **260** 166
[107] Reynolds D C, Litton C W, Look D C, Hoelscher J E, Claflin B, Collins T C, Nause J and Nemeth B 2004 *J. Appl. Phys.* **95** 4802
[108] Nause J and Nemeth B 2005 *Semicond. Sci. Technol.* **20** S45
[109] Ardakani H K 1996 *Thin Sol. Films* **287** 280
[110] Quaranta F, Valentini A, Rizzi F R and Casamassima G 1993 *J. Appl. Phys.* **74** 244
[111] Tuomisto F, Saarinen K, Look D C and Farlow G C 2005 *Phys. Rev.* B **72** 085206
[112] Look D C, Hemsky J W and Sizelove J R 1999 *Phys. Rev. Lett.* **82** 2552
[113] Ohtomo A, Kawasaki M, Koida T, Masubuchi K, Koinuma H, Sakurai Y, Yoshida Y, Yasuda T and Segawa Y 1998 *Appl. Phys. Lett.* **71** 2466
[114] Ohtomo A, Tamura K, Kawasaki M, Makino T, Segawa Y, Tang Z K, Wong L G K, Matsumoto Y and Koinuma H 2000 *Appl. Phys. Lett.* **77** 2204
[115] Jin Y, Zhang B, Shuming Y, Wang Y, Chen J, Zhang H, Huang C, Cao C, Cao H and Chang R P H 2001 *Solid State Commun.* **119** 409
[116] Makino T, Segawa Y, Ohtomo A, Tamura K and Koinuma H 2001 *Appl. Phys. Lett.* **78** 1237
[117] Gruber Th and Kirchner C and Kling R and Reuss F and Waag A 2004 *Appl. Phys. Lett.* **84** 5359
[118] Gruber Th, Kirchner C, Kling R, Reuss F, Waag A, Bertram F, Forster D, Christen J and Schreck M 2003 *Appl. Phys. Lett.* **83** 3290
[119] Nomura K, Hiromichi O, Takagi A, Kamiya T, Hirano M and Hosono H 2004 *Nature* **432** 488
[120] Hoffman R L, Norris B J and Wager J F 2003 *Appl. Phys. Lett.* **82** 733
[121] Hoffman R L 2004 *J. Appl. Phys.* **95** 5813
[122] Heiland H, Mollwo E and Stoeckmann F 1959 *Solid State Phys.* **8** 191
[123] Scharowski E 1953 *Z. Phys.* **135** 138
[124] Dodson E M and Savage J A 1968 *J. Mater. Sci.* **3** 19
[125] Helbig R 1972 *J. Cryst. Growth* **15** 25
[126] Laudise R A and Ballmann A A 1960 *J. Phys. Chem.* **64** 688
Laudise R A, Kolb E D and Caporaso A J 1964 *J. Am. Ceram. Soc.* **47** 9
[127] Tuomisto F, Ranki V, Saarinen K and Look D C 2003 *Phys. Rev. Lett.* **91** 205502
[128] Ryu Y R, Zhu S, Budai J D, Chandrasekhar H R, Miceli P F and White H W 2000 *J. Appl. Phys.* **88** 201
[129] Tsukazaki A, Ohtomo A, Yoshida S, Kawasaki M, Chia C H, Makino T, Segawa Y, Koida T, Chichibu S F and Koinuma H 2003 *Appl. Phys. Lett.* **83** 2784
[130] Guo X-L, Tabata H and Kawai T 2001 *J. Cryst. Growth* **223** 135
[131] Hwang D-K, Bang K-H, Jeong M-C and Myoung J 2003 *J. Cryst. Growth* **254** 449
[132] Ji Z G, Liu K, Yang C, Fan R X and Ye Z Z 2003 *J. Cryst. Growth* **253** 246
Ye Z Z, Lu J G, Chen H H, Zhang Y Z, Wang L, Zhao B H and Huang J Y 2003 *J. Cryst. Growth* **253** 258
[133] Wang C, Ji Z G, Liu K, Xiang Y and Ye Z Z 2003 *J. Cryst. Growth* **259** 279
[134] Wang J *et al* 2003 *J. Cryst. Growth* **255** 293
[135] Bian J M, Li X M, Zhang C Y, Chen L D and Yao Q 2004 *Appl. Phys. Lett.* **84** 3783
[136] Ye Z Z, Ge F, Lu J, Zhang Z, Zhu L, Zhao B and Huang J 2004 *J. Cryst. Growth* **265** 127
Xu W, Ye Z Z, Zhou T, Zhao B, Zhu L and Huang J 2004 *J. Cryst. Growth* **265** 133
[137] Graubner S, Neumann C, Volbers N, Meyer B K, Bläsing J and Krost A 2007 *Appl. Phys. Lett.* **90** 042103
[138] Neumann C, Lautenschlager S, Graubner S, Sann J, Volbers N, Meyer B K, Bläsing J, Krost A, Bertram F and Christen J 2007 *Phys. Status Solidi* b **244** 1451
[139] Shaer A E, Bakin A, Mofor A C, Bläsing J, Krost A, Stoimenos J, Pecz B, Kreye M and Waag A 2007 *Appl. Phys. A: Mater. Sci. Eng.* **88** 57
[140] Triboulet R, Perrière J 2003 *Prog. Cryst. Growth Charact. Mater.* **47** 65
[141] Pearton S J, Norton D P, Ip K and Heo Y W 2004 *J. Vac. Sci. Technol.* B **22** 932
[142] Grundmann M, von wenckstern H, Pickenhain R, Weinhold S, Chengnui B and Breinstein O 2005 *Zinc Oxide—A Material for Micro- and Optoelectronic*

Applications ed N H Nickel and E Terukov (Netherlands: Springer) chapter 5
[143] Gupta V and Sreenivas K 2006 *Zinc Oxide Bulk, Thin Films, and Nanostructures* ed C Jagadish and S J Pearton (New York: Elsevier) chapter 4
[144] Coutts T J, Li X, Barnes T M, Keyes B M, Craig L P, Asher S E, Zhang S B and Wei S-H 2006 *Zinc Oxide Bulk, Thin Films, and Nanostructures* ed C Jagadish and S J Pearton (New York: Elsevier) chapter 3
[145] Johnson M A L, Fujita S, Rowland W H, Hughes W C, Cook J W and Schetzina J F 1996 *J. Electron. Mater.* **25** 855
[146] Zhao J L, Li X M, Krtschil A, Krost A, Yu W D, Zhang Y W, Gu Y F and Gao X D 2007 *Appl. Phys. Lett.* **90** 062118
[147] Kittel C 2005 *Introduction to Solid State Physics* 8th edn (New York: Wiley)
[148] Van de Walle C G and Neugebauer J 2004 *J. Appl. Phys.* **95** 3851
[149] Laks D B, Van de Walle C G, Neumark G F, Blöchl P E and Pantelides S T 1992 *Phys. Rev.* B **45** 10965
[150] Dean J A 1992 *Lange's Handbook of Chemistry* 14th edn (New York: McGraw-Hill)
[151] Makov G and Payne M C 1995 *Phys. Rev.* B **51** 4014
[152] Shim J, Lee E-K, Lee Y-J and Nieminen R M 2005 *Phys. Rev.* B **71** 035206
[153] Freysoldt C, Neugebauer J and Van de Walle C G 2009 *Phys. Rev. Lett.* **102** 016402
[154] Mooney P M 1999 *Identification of Defects in Semiconductors* ed M Stavola *Semiconductors and Semimetals* vol 51B (San Diego: Academic) p 93
[155] Neumann G 1981 *Current Topics in Materials Science* vol 7 ed E Kaldis (Amsterdam: North Holland)
[156] Glicksman M E 2000 *Diffusion in Solids: Field Theory, Solid-State Principles and Applications* (New York: Wiley)
[157] Nichols C S, Van de Walle C G and Pantelides S T 1989 *Phys. Rev. Lett.* **62** 1049
[158] Limpijumnong S and Van de Walle C G 2004 *Phys. Rev.* B **69** 035207
[159] Janotti A, Krčmar M, Fu C L and Reed R 2004 *Phys. Rev. Lett.* **92** 085901
[160] Watkins G D 1976 *Radiation Effects in Semiconductors* ed N B Urli and J W Corbett *IOP Conf. Proc.* No 31 (London: Institute of Physics) p 95
[161] Chow K H and Watkins G D 1998 *Phys. Rev. Lett.* **81** 2084
[162] Gorelkinskii Yu V and Watkins G D 2004 *Phys. Rev.* B **69** 115212
[163] Coskun C, Look D C, Farlow G C and Sizelove J R 2004 *Semicond. Sci. Technol.* **19** 752
[164] Vineyard G H 1957 *J. Phys. Chem. Solids* **3** 121
[165] Janotti A, Segev D and Van de Walle C G 2006 *Phys. Rev.* B **74** 045202
[166] Smith J M and Vehse W E 1970 *Phys. Lett.* A **31** 147
[167] Gonzalez C, Galland D and Herve A 1975 *Phys. Status Solidi* b **72** 309
[168] Soriano V and Galland D 1976 *Phys. Status Solidi* b **77** 739
[169] Schneider J and Räuber A 1961 *Z. Naturf.* a **16** 712
[170] Kasai P H 1963 *Phys. Rev.* **130** 989
[171] Müller K A and Schneider J 1963 *Phys. Lett.* **4** 288
[172] Geisler C H and Simmons G L 1964 *Phys. Lett.* **11** 111
[173] Lal R B and Arnett G M 1966 *J. Phys. Soc. Japan* **21** 2743
[174] Sancier K M 1970 *Surface Sci.* **21** 1
[175] Born G K, Hofstaetter A B, Scharmann A O, Arnett G M, Kroes R L and Wegner U E 1971 *Phys. Status Solidi* a **4** 675
[176] Mookherji T 1972 *Phys. Status Solidi* a **13** 293
[177] Pöppl A and Völkel G 1989 *Phys. Status Solidi* a **115** 247
[178] Pöppl A and Völkel G 1990 *Phys. Status Solidi* a **121** 195
[179] Vanheusden K, Seager C H, Warren W L, Trallant D R, Caruso J, Hampden-Smith M J and Kodas T T 1997 *J. Lumin.* **75** 11
[180] Carlos W E, Glaser E R and Look D C 2001 *Physica* B **308–310** 976
[181] Garces N Y, Giles N C, Halliburton L E, Cantwell G, Eason D B, Reynolds D C and Look D C 2002 *Appl. Phys. Lett.* **80** 1334
[182] Taylor A L, Filipovich G and Lindeberg G K 1970 *Solid State Commun.* **8** 1359
[183] Tarkpea K, Ots A and Nikitenko V 1994 *J. Phys. Chem. Solids* **55** 1353
[184] Neumann G 1981 *Current Topics in Materials Science* vol 7 ed E Kaldis (Amsterdam: North Holland) p 269
[185] Leiter F H, Alves H R, Hofstaetter A, Hofmann D M and Meyer B K 2001 *Phys. Status Solidi* b **226** R4
[186] Leiter F, Alves H, Pfisterer D, Romanov N G, Hofmann D M and Meyer B K 2003 *Physica* B **340–342** 201
[187] Hofmann D M, Pfisterer D, Sann J, Meyer B K, Tena-Zaera R, Munoz-Sanjose V, Frank T and Pensl G 2007 *Appl. Phys. A Mater. Sci. Proc.* **88** 147
[188] Evans S M, Giles N C, Halliburton L E and Kappers L A 2008 *J. Appl. Phys.* **103** 043710
[189] Galland D and Herve A 1974 *Solid State Commun.* **14** 953
[190] Tuomisto F, Saarinen K and Look D C 2004 *Phys. Status Solidi* **201** 2219
[191] Galland D and Herve A 1970 *Phys. Lett.* A **33** 1
[192] Lauer R B 1973 *J. Phys. Chem. Solids* **34** 249
[193] Reynolds D C, Look D C, Jogai B and Morkoç H 1997 *Solid State Commun.* **101** 643
[194] Reynolds D C, Look D C, Jogai B, Van Nostrand J E, Jones R and Jenny J 1998 *Solid State Commun.* **106** 701
[195] Lin B, Fu Z and Jia Y 2001 *Appl. Phys. Lett.* **79** 943
[196] Dingle R 1969 *Phys. Rev. Lett.* **23** 579
[197] Mishra K C, Schmidt P C, Johnson K H, DeBoer B G, Berkowitz J K and Dale E A 1990 *Phys. Rev.* B **42** 1423
[198] Neugebauer J and Van de Walle C G 1996 *Appl. Phys. Lett.* **69** 503
[199] Kröger F A and Vink H J 1954 *J. Chem. Phys.* **22** 250
[200] Studenikin S A, Golego N and Cocivera M 1998 *J. Appl. Phys.* **84** 2287
[201] Vanheusden K, Warren W L, Seager C H, Trallant D R and Voigt J A 1996 *J. Appl. Phys.* **79** 7983
[202] Sekiguchi T, Ohashi N and Terada Y 1997 *Japan. J. Appl. Phys.* **36** L289
[203] Lavrov E V, Weber J, Börrnert F, Van de Walle C G and Helbig R 2002 *Phys. Rev.* B **66** 165205
[204] Look D C, Farlow G C, Reunchan P, Limpijumnong S, Zhang S B and Nordlund K 2005 *Phys. Rev. Lett.* **95** 225502
[205] Tomlins G W, Routbort J L and Mason T O 2000 *J. Appl. Phys.* **87** 117
[206] Limpijumnong S, Li X, Wei S-H and Zhang S B 2005 *Appl. Phys. Lett.* **86** 211910
[207] Hu J and Gordon R G 1992 *J. Appl. Phys.* **71** 880
[208] Ko H J, Chen Y F, Hong S K, Wenisch H, Yao T and Look D C 2000 *Appl. Phys. Lett.* **77** 3761
[209] Block D, Hervé A and Cox R T 1982 *Phys. Rev.* B **25** 6049
[210] Hu J and Gordon R G 1992 *J. Appl. Phys.* **72** 5381
[211] Hu J and Gordon R G 1993 *Mater. Res. Soc. Symp. Proc.* **283** 891
[212] Wenas W W *et al* 1991 *J. Appl. Phys.* **70** 7119
[213] Hu J and Gordon R G 1991 *Solar Cells* **30** 437
[214] Pankove J I and Johnson N M (ed) 1991 *Hydrogen in Semiconductors*, *Semiconductors and Semimetals* vol 34 (Boston: Academic)
[215] Van de Walle C G and Johnson N M 1998 *Gallium Nitride (GaN) II* (*Semiconductors and Semimetals* vol 57)

ed J I Pankove and T D Moustakas (Boston: Academic) p 157
[216] Van de Walle C G and Neugebauer J 2006 *Annu. Rev. Mater. Res.* **36** 179
[217] Myong S Y *et al* 1997 *Japan. J. Appl. Phys.* **36** L1078
[218] Mollwo E 1954 *Z. Phys.* **138** 478
[219] Lander J J 1957 *J. Phys. Chem. Solids* **3** 87
[220] Baik S J, Jang J H, Lee C H, Cho W Y and Lim K S 1997 *Appl. Phys. Lett.* **70** 3516
[221] Kohiki S *et al* 1994 *Appl. Phys. Lett.* **64** 2876
[222] Yaron G, Levy J, Many A, Goldstein Y, Weisz S Z and Resto O 1986 *J. Phys. Chem. Solids* **47** 401
[223] Lander J J 1960 *J. Phys. Chem. Solids* **15** 324
[224] Kolb E D and Laudise R A 1965 *J. Electrochem. Soc.* **48** 344
[225] Schirmer O F 1968 *J. Phys. Chem. Solids* **29** 1407
[226] Meyer B K, Stehr J, Hofstaetter A, Volbers N, Zeuner A and Sann J 2007 *Appl. Phys.* A **88** 119
[227] Lee E-C and Chang K J 2006 *Physica* B **376** 707
[228] Zeng Y J, Ye Z Z, Lu J G, Xu W Z, Zhu L P, Zhao B H and Limpijumnong S 2006 *Appl. Phys. Lett.* **89** 042106
[229] Laegsgaard J and Stokbro K 2001 *Phys. Rev. Lett.* **86** 2834
[230] Müller G 1976 *Phys. Status Solidi* b **76** 525
[231] Dietz R E, Kamimura H, Sturge M D and Yariv A 1963 *Phys. Rev.* **132** 1559
[232] Thonke K, Gruber Th, Teofilov N, Schöfelder R, Kerwien N, Waag A and Sauer R 2001 *Physica* B **308** 945
[233] Guo X-L, Choi J-H, Tabata H and Kawai T 2001 *Japan. J. Appl. Phys.* **40** L177
[234] Meyer B K *et al* 2004 *Phys. Status Solidi* b **241** 231
[235] Lautenschlaeger S, Eisermann S, Meyer B K, Callison G, Wagner M R and Hoffmann A 2009 *Phys. Status Solidi RRL* **3** 16
[236] Fang Z-Q, Claflin B, Look D C, Lei Kerr L and Li X 2007 *J. Appl. Phys.* **102** 023714
[237] Fons P, Tampo H, Kolobov A V, Ohkubo M, Niki S, Tominaga J, Carboni R, Boscherini F and Friedrich S 2006 *Phys. Rev. Lett.* **96** 045504
[238] Van de Walle C G, Laks D B, Neumark G F and Pantelides S T 1993 *Phys. Rev.* B **47** 9425
[239] Kobayashi A, Sankey O F and Dow J D 1983 *Phys. Rev.* B **28** 946
[240] Chadi D J 1991 *Appl. Phys. Lett.* **59** 3589
[241] Lee W J, Kang J and Chang K J 2006 *Phys. Rev.* B **73** 024117
[242] Neugebauer J and Van de Walle C G 1996 *Appl. Phys. Lett.* **68** 1829
[243] Yamamoto T and Katayama-Yoshida H 1999 *Japan. J. Appl. Phys.* **38** L166
[244] Van de Walle C G, Limpijumnong S and Neugebauer J 2001 *Phys. Rev.* B **63** 245205
[245] Zhang S B, Wei S-H and Yan Y 2001 *Physica* B **302–303** 135
[246] Yan Y, Zhang S B and Pantelides S T 2001 *Phys. Rev. Lett.* **86** 5723
[247] Yan Y, Li Ji, Wei S-H and Al-Jassim M M 2007 *Phys. Rev. Lett.* **98** 135506
[248] Wassner T A, Laumer B, Maier S, Laufer A, Meyer B K, Stutzmann M and Eickhoff M 2009 *J. Appl. Phys.* **105** 023505
[249] Lautenschlaeger S, Sann J, Klar P J, Piechotka M and Meyer B K 2009 *Phys. Status Solidi* b **246** 383
[250] Kim Y-I, Page K and Seshadri R 2007 *Appl. Phys. Lett.* **90** 101904
[251] Kim Y-I, Page K, Limarga A M, Clarke D R and Seshadri R 2007 *Phys. Rev.* B **76** 115204
[252] Kim Y-I, Cadars S, Shayib R, Proffen Th, Feigerle C S, Chmelka B F and Seshadri R 2008 *Phys. Rev.* B **78** 195205
[253] Tsukazaki A, Ohtomo A, Kita T, Ohno Y, Ohno H and Kawasaki M 2007 *Science* **315** 1388
[254] Tampo H *et al* 2008 *Appl. Phys. Lett.* **93** 202104
[255] Nishimoto Y *et al* 2008 *Appl. Phys. Express* **1** 091202
[256] Chauveau J-M, Morhain C, Teisseire M, Laugt M, Deparis C, Zuniga-Perez J and Vinter B 2009 *Microelectron. J.* **40** 512
[257] Van de Walle C G 1989 *Phys. Rev.* B **39** 1871
[258] Janotti A and Van de Walle C G 2007 *Phys. Rev.* B **75** 121201
[259] Van de Walle C G and Martin R 1989 *Phys. Rev. Lett.* **62** 2028
[260] Van de Walle C G and Neugebauer J 1997 *Appl. Phys. Lett.* **70** 2577
[261] Neugebauer J and Van de Walle C G 1994 *Phys. Rev.* B **50** 8067
[262] Neugebauer J and Van de Walle C G 1996 *Festkörperprobleme/Advances in Solid State Physics* vol 35 ed R Helbig (Vieweg: Braunschweig/Wiesbaden) p 25