Chapter 1
Numerical Simulations of Defective Structures: The Nature of Oxygen Vacancy in Non-reducible (MgO, SiO₂, ZrO₂) and Reducible (TiO₂, NiO, WO₃) Oxides

Gianfranco Pacchioni

Abstract The nature of a common defect in oxide materials, the oxygen vacancy, is analyzed from a theoretical point of view based on first principles density functional theory calculations. Different oxides with non-reducible and reducible character are compared to show that the electronic structure and the properties of this defect are strongly related to the kind of chemical bond present in the extended material. This also results in different formation energies of the defect, an aspect that directly affects the surface chemistry of these systems.

1.1 Introduction: The Role of Defects in Oxide Materials

The functioning of several advanced devices is largely related to the presence of faults or defects in the structure of the material used for their construction [1, 2]. Material and surface properties are strongly dependent on the nature, concentration, and arrangement of defects of various dimensionality like pores, voids, grain boundaries, dislocations, impurities, and missing atoms (vacancies). One of the fields where the nature of defects has been used to produce new functionalities is microelectronics, in which the properties of Si and the III–V (e.g. GaAs) semiconductors and their use in practical devices depend on the control of dopants, interface states, etc. Regarding mechanical properties, the strength of a metal is largely related to the presence of dislocations; if one considers optical properties, the color of a gem is due to the presence of very small amounts of transition metal
atom impurities embedded in an oxide matrix. In connection with transport properties, the superconducting behavior of some cuprates is closely connected to the level of oxygen vacancies. Given their refractory nature, oxides are often used at elevated temperatures and are exposed to strongly oxidising or reducing atmospheres, with consequent exchange of matter with the gas-phase and intrinsic generation of defects. In chemistry and heterogeneous catalysis, surface defects are extremely important as they largely determine corrosion resistance, molecular adsorption, and conversion of one chemical species into another.

The study of the chemical, electrical and optical properties of defects in oxide materials plays an crucial role in material science, solid state chemistry and solid state physics. Defects play a dominant role also in surface and interface properties, and in nanotechnology. The control of the amount, nature and characteristics of defects in solids is at the basis of defect engineering, a discipline aimed at using defects in a beneficial way to tune material properties in a desired manner or to generate new material behaviors (see also Chap. 14). There are numerous fields where the inclusion of specific defects has improved the performances of a given material or device. Perhaps, one of the topics where more efforts has been directed in the last decades is in the search of new photo-active materials able to absorb visible light and to efficiently generate electron-hole pairs that can turn an inactive oxide into a photo-catalytic material to be used in environmental catalysis or for energy production [3]. Nitrogen doped TiO$_2$ is one particular system that has attracted a lot of interest [4–7]. The N-dopant introduces new states in the gap which lower the excitation energy and improve the photo-absorption of the material. Unfortunately, this also results in a much easier formation of oxygen vacancies hence in the creation of sites that help the electron-hole recombination, a process detrimental for the photo-catalytic properties [6, 7]. Thus, the presence of heteroatoms can induce profound consequences on the stability of the oxide, on the nature of its surface, and on its capability to release oxygen. In the second half of last century a huge effort has been made to identify, characterise, and control defects in silicon dioxide, a key material for optical fibers and metal-oxide-semiconductor field-effect transistors that form the basis of the present revolution in information and telecommunication technology. While in the past the attention has been almost exclusively on bulk defects, in the last two decades it has moved to defects formed at the surfaces of oxides [8].

The increase in interest in the control of defects in materials has grown in parallel with the development of new theoretical methods, algorithms and protocols to simulate and predict material and surface properties. This, together with the exponential growth of computing power, has made the computer simulation of defects in bulk materials and their surfaces an essential tool to complement and reinforce the experimental information. Still, the description of defects at oxide surfaces is not yet fully satisfactory. While some oxides (e.g. MgO, SiO$_2$, TiO$_2$, CeO$_2$) have been studied in detail, and the level of understanding of their bulk and surface defectivity can be considered good, other oxides are much less understood. This is the case in particular for complex transition metal oxides, for ternary compounds, for oxide nanostructures (nanoparticles, thin films), etc.
In this chapter we are interested in some general features of defects in oxide materials (both bulk and surface) and in the problems related to their description with modern electronic structure theory, in particular descriptions based on the density functional theory (DFT) formalism. To illustrate these problems, take the determination of reaction energies from first principles calculations [9]. Several industrial catalysts are designed to perform oxidation or oxidative dehydrogenation reactions. The catalysts used in these processes are metal oxides and the mechanism which is normally observed is that introduced by Mars and van Krevelen [10], where an organic compound reacts with specific surface sites (e.g. more reactive oxygen atoms at low-coordinated sites) and an oxygen atom is transferred from the surface to the organic molecule with consequent formation of a vacancy and modification of the stoichiometry of the material. In an oxygen atmosphere, O₂ molecules interact with the surface, dissociate into oxygen atoms that can diffuse and eventually fill the vacancy created during the oxidative process thus restoring the original stoichiometry and composition of the catalyst. This mechanism has been identified in several catalytic reactions by using isotopically labeled oxygen to prove that the atom incorporated in the organic molecule does not come from the gas phase but rather from the oxide surface [11]. The cost of forming the oxygen vacancy by removing the oxygen atom from the surface is a key parameter that determines both the kinetics and the thermodynamics of the reaction. Contrary to what one could expect, this quantity is rather difficult to determine accurately, as its value greatly depends on the details of the calculations, on the specific exchange-correlation functional used, on the description of the defect center formed (localised or delocalised nature of trapped electrons), etc. Since uncertainties of a few kcal/mole in an energy barrier translate into kinetic constants that differ by orders of magnitude, it is clear that the development of accurate methods for the determination of defects formation energies is highly desirable yet not easily attainable.

On the other hand, theory can be extremely useful in predicting trends. Given the role that oxygen release from an oxide surface has on surface chemistry and catalysis, several attempts have been made to modify the cost of oxygen removal from an oxide by doping the material with heteroatoms [12]. Specific efforts have been made to design better catalytic materials, in particular for oxidation reactions [12–14]. This topic has been extensively covered in a recent review by McFarland and Metiu [15]. The replacement of a metal cation in an oxide of MₓOₙ formula with another dopant X opens several possibilities to selectively modify the electronic structure of the oxide. If the doping heteroatom X has the same valency of the metal cation M, the changes are mostly related to the different size of the two cations and to the different strengths of the M-O and X-O bonds. However, if the X element has a different number of valence electrons several possibilities exist to compensate the charge. This opens a complex scenario of defects that can more easily form and appear in the material, also as a function of the preparation conditions.

A reaction that for a long time has been considered a typical example of catalysts based on metal doping of oxides is the coupling of methane, CH₄, to give ethane, C₂H₆, and other C₂ hydrocarbons by Li-doped MgO [16]. The mechanism which
has been accepted until recently is that the Li impurity creates O radicals whose presence can be detected by electron paramagnetic resonance (EPR) experiments [17]. The O radicals are very reactive and help the extraction of an H atom from methane to form a methyl radical, which then further reacts to form ethane and ethene. As this mechanism seems both sufficiently simple and reasonable, it has been accepted by the catalysis community for a few decades. However, the problem has been recently reconsidered with a large effort based on several experimental and computational studies leading to some surprising results [18]. In fact, it has been found that no radical species are actually detected in the working catalyst and that the major role of Li is to induce a structural modification of the oxide [18]. This example shows in a very clear way how difficult it is to identify and characterize defects in oxides. The difficulty is mainly related to the very low concentrations, sometimes sub-part-per-million level, of the defects. This requires the use of very sensitive experimental techniques, and often the combination of more than a single technique.

First principles simulations can be of great help for the identification of potentially beneficial or detrimental dopants and for the design of new catalytic materials. DFT, in one of its several versions, is presently the most popular approach to determine the electronic structure and properties of materials. DFT has almost completely replaced more “classical” solutions to solve the Schrödinger equation based on the determination of the wave function of the system. These are the so-called wave-function based methods, like the Hartree-Fock (HF) and the post-HF approaches where correlation effects are included with specific, often computationally very demanding, procedures. Still these methods play a very important role, for instance in the determination of optical excitations in solids [19]. Improvements in the formulation of exchange-correlation functionals used in DFT are continuously appearing in the literature, but still there are open questions about the reliability of DFT to study some special material properties. Particularly challenging is the determination of the electronic structure of transition metal oxides and other highly correlated systems. Not surprisingly, similar kinds of problems arise when one deals with the description of bulk and surface defects in oxides.

1.2 Treating Defects in Solids: Periodic Models and Local Modes

1.2.1 Periodic Models

The electronic structure of solids and surfaces is usually described in terms of band structure. To this end, a unit cell containing a given number of atoms is periodically repeated in three dimensions to account for the “infinite” nature of the crystalline solid and the Schrödinger equation is solved for the atoms in the unit cell subject to periodic boundary conditions [20]. This approach can also be extended to the study
of surface or of bulk defects by means of the supercell approach in which an artificial periodic structure is created where the defect is translationally reproduced in correspondence to a given super-lattice of the host. This procedure allows the use of efficient computer programs designed for the treatment of periodic systems and has indeed been followed by several authors to study defects using either DFT and plane waves approaches [21–23] or HF based methods with localised atomic orbitals [24, 25]. The quality of the basis set when using plane waves approaches is controlled by one parameter, the kinetic energy cutoff, but for a given cutoff the number of plane waves depends on the size of the simulation cell.

The presence of the defect in the unit cell, however, results in a periodic repetition in the three directions of space, hence modelling high defect density. The only way to reduce the defect concentration is to increase the size of the unit cell, a solution which implies a large computational cost. Nowadays, periodic calculations for supercells containing a few hundreds atoms are routinely done. Even for large supercells containing ≈200 atoms, however, the defect density may still be too large. The supercell approach is therefore based on the assumption that the defects do not interact appreciably except when they are very close to each other, so that rapid convergence is achieved by increasing the size of the supercell. With charged defects (e.g. ions included in the bulk material, electron traps, etc.), quite common in insulators, the supercell approach is feasible but more problematic due to a divergent electrostatic energy. Conventionally, a charge-neutrality condition is imposed, which implicitly introduces an artificial homogeneous and oppositely charged (“jellium”) background [26]. Computation of bulk and surface defect formation energies in the dilute limit are computationally demanding because large unit cells are required, and the true limit remains practically inaccessible.

1.2.2 Local Cluster Models

An alternative to the periodic band structure methods for studying solids is the cluster approach [27, 28]. Here one explicitly considers only a finite number of atoms to describe a part of the surface while the rest is treated in a more or less simplified way (embedding, see below). The main conceptual difference is that in the cluster approach one uses molecular orbitals, MO, instead of delocalised functions (e.g. plane waves). The description of the electronic properties is thus done in terms of local orbitals, allowing one to treat problems in solids with the typical language of chemistry, the language of orbitals. This is particularly useful when dealing with surface problems and with the reactivity of a solid surface. In fact, the interaction of gas-phase molecules with a solid surface can be described in exactly the same way as the interaction of two molecules. Of course, the cluster model is also not free from limitations. The most serious one is that the effect of the surrounding is often taken into account in a more or less approximate way, thus leading to some uncertainties in the absolute values of the computed quantities. It is also possible that some properties are described differently depending on the size of
the cluster used. It is therefore necessary to check the results against cluster size and shape. The advantages are (1) a very low concentration of defects is considered so that no mutual defect-defect interaction is present in the model and (2) accurate theoretical methods derived from quantum chemistry can be applied (like configuration interaction, Møller–Plesset (MP) perturbation theory and coupled-cluster (CC) methods). This is an important advantage which should not be underestimated. In fact, in this way it is possible to explicitly include correlation effects in the calculations and to treat exactly the non-local exchange as in the HF formalism. This second aspect can be particularly important for the description of magnetic systems, radical species, optical transitions or defects with localised holes or electrons [29].

Therefore, cluster calculations represent an alternative way of describing localised bonds at surfaces as well as defects in ionic crystals. The problem is how to introduce the effect of the rest of the crystal. Completely different strategies can be adopted to “embed” clusters of largely covalent oxides, like SiO₂, or of very ionic oxides, like MgO. In SiO₂ and related materials the cluster dangling bonds are usually saturated by H atoms [27]. The saturation of the dangling bonds with H atoms is an important aspect of the embedding, but not the only one. In fact, in this way one neglects the crystalline Madelung field. This term is quite important for the description of solid surfaces with more pronounced ionic character, like MgO, but also of semi-covalent materials like silica. Much more complex is the treatment of systems which have a mixed ionic-covalent nature, or easily reducible materials like transition metal oxides. Here, the ions at the cluster boundary tend to act as charge traps and behave, themselves, as defects. A simple way to embed clusters of reducible oxides and to take into account solid state effects does not exists yet.

1.2.3 Embedding Schemes

The very ionic nature of MgO and other ionic oxides implies that the Madelung potential is explicitly included. Indeed, several properties of ionic crystals are incorrectly described if the long range Coulomb interactions are not taken into account [30]. A crude approach is to surround the cluster of ions by a large array of point charges (PC) to reproduce the Madelung field of the host at the central region of the cluster [31]. However, the PC’s polarise the oxide anions at the cluster border and cause an incorrect behavior of the electrostatic potential [32]. The problem can be eliminated by placing an effective core potential, ECP, at the position of the positive PC’s around the cluster, representing the finite size of the cation core [33]. No basis functions are associated with the ECP which accounts for the Pauli or exchange repulsion of the O²⁻ valence electrons with the surrounding. This is a simplified approach to the more rigorous ab initio model potential (AIMP) method [34, 35] but is computationally simple and reliable. In the AIMP approach the grid of bare charges is replaced by a grid of AIMP’s which account not only for the long-range Coulomb interaction but also for the quantum mechanical short-range
requirements of exchange and orthogonality without explicitly introducing extra electrons in the model.

The addition of the ECP’s to the cluster gives a better representation of the electrostatic potential and, hence, the electrostatic contribution to the surface bonding. What is still missing from this simplified approach is the polarisation of the host crystal induced by an adsorbed species or by the presence of a defect. This effect can be particularly important for charged adsorbates or defects. The polarisation, $E_{\text{pol}}$, induced by a charge on the surrounding lattice can be estimated by means of the classical Born formula [36]:

$$E_{\text{pol}} = -\left(1 - 1/\varepsilon\right) q^2 / 2R$$  \hspace{1cm} (1.1)

where $\varepsilon$ is the dielectric constant of the material, $q$ is the absolute value of the charge and $R$ is the radius of the spherical cavity where the charge is distributed. Since a certain degree of ambiguity remains in the definition of $R$, this correction is only qualitative.

A treatment based on a mixed quantum-classical approach has been proposed [37]. The ionic crystal is represented by a large finite nano-cluster, which is divided into two regions: region I, centered on the defect site, and the rest—region II. Region I includes a quantum-mechanically treated cluster (QM cluster) surrounded by interface ions and a region of classical shell model ions [38]. The remaining part of the nano-cluster is represented by point charges. The classical ions, both shell-model and PC’s, interact among themselves via classical interatomic potentials. All quantum-mechanical, interface and classical ions (both cores and shells) in region I are allowed to relax simultaneously in the course of geometry optimisation. Ions outside region I remain fixed and provide accurate electrostatic potentials within region I. This approach allows one to take into account the defect-induced lattice polarisation of a very large crystal region. In bulk calculations it can be extended to infinity outside region I using a polarisable continuum model and the Mott-Littleton approach. The interaction between the QM atoms and classical atoms in region I and II is also included and represented using short-range classical potentials. In this way it has been possible to include both geometric and electronic relaxations in a region of several Å in diameter at a reasonable computational cost [37].

The approach described above can be considered as a special case of a more general kind of approaches called quantum mechanics/molecular mechanics (QM/MM) methods. These represent a good compromise between accuracy and computational cost [39]. Among these methods, particular popular are the integrated molecular orbital molecular mechanics (IMOMM) [40] or ONIOM [41] methods. All these approaches have the particularity of designing a small quantum region treated accurately and a surrounding region treated with either force fields or less accurate quantum chemical methods. The main problem consists of defining the interface between both regions, which is made by saturating the dangling bonds with hydrogen atoms. The IMOMM and ONIOM methods permit one to handle large models and to include all important effects of the crystal.
1.3 Problems of DFT in Describing Defects in Insulators: Some Instructive Examples

A well-known problem in DFT solutions of the Schrödinger equation is the self-interaction [42, 43]. The HF energy contains no self-interaction contributions because the self-interaction part of the Coulomb energy cancels exactly that of the exchange part. This requirement cannot be satisfied in DFT without special effort. Because of the self-interaction, in DFT the unpaired electrons tend to delocalize over many atoms to reduce the Coulomb repulsion. This problem severely affects the description of complex systems like narrow band oxides [44], and has also been recognised for molecular systems [45]. Standard DFT describes a magnetic insulator like NiO as a metal, giving a completely wrong picture. The same problem occurs when it comes to the description of defects in insulators or wide gap semiconductors. A case which shows dramatically the weakness of the standard DFT approach is that of an Al impurity in bulk SiO2. Al-doped SiO2 has been studied in great detail, both theoretically and experimentally since the 1950s [46–50]. Early EPR experiments showed that the defect center, corresponding to an Al atom substituting a four-coordinated Si in the lattice, [AlO4]0, contains a hole trapped in a non-bonding 2p orbital of an O atom adjacent to Al [46, 48]. The fact that the hole is localised, at least at low temperature, is shown by EPR in particular by the hyperfine interaction of the electron spin with the 27Al, 29Si and 17O nuclides [48]. Above room temperature, the hole hops rapidly among all four adjacent O atoms and becomes delocalized. Cluster calculations performed at the HF level in the 1980s confirmed this model, and the associated elongation of a single Al-O bond [49, 50]. With the advent of DFT, this model was questioned by new studies based on supercell calculations with the inclusion of boundary conditions (DFT at the LDA or GGA level) [51–53]. At the DFT level the hole in [AlO4]0 is completely delocalized over the four O neighbors [51–53], even at zero Kelvin. The different physical picture emerging from cluster HF and supercell DFT calculations is not related to the model used (cluster or periodic) but is due to the non-exact treatment of the exchange term in DFT [29]. This problem, which is now well known, can be partly solved by using the so-called hybrid functionals where a given portion of HF exchange is mixed in with the DFT exchange functional. This approach was originally proposed by Becke and implemented in the B3LYP hybrid functional [54, 55], and later expanded in a new series of theoretically more well grounded hybrid functionals (PBE0, HSE06, etc.) which are particularly efficient for solid state problems as shown by some recent studies on rare-earth oxides [56–58]. The [AlO4]0 center, however, is a particularly complex problem and even B3LYP and other hybrid functionals fail to provide the correct description. In order to get a fully localised hole on a single O atom one has to increase the amount of HF exchange (20 % in B3LYP) to at least 50 % [59–61]. The problem has been recently reconsidered by Gillen and Robertson who used a screened-exchange hybrid functional to successfully describe spin localisation in this defect [62].
The problem of an Al impurity in SiO$_2$ has also been addressed using the so-called DFT+U approach [63]. This method was developed to account for correlation effects arising from the strong on-site Coulomb repulsion and exchange interactions which determine the structure of Mott-Hubbard insulators. In DFT+U approaches (either in the LDA or GGA variant) one chooses a set of atomic-like orbitals which are treated with a new Hamiltonian [64, 65]. In the approach of Dudarev et al. [66] the latter depends on the difference $U_{\text{eff}} = U - J$, where $U$ is a parameter which describes the energy increase for an extra electron on a particular site and $J$ a second parameter which represents the screened exchange energy. This approach corrects most of the inadequacies connected to the treatment of localized states, but suffers from the dependence of the results on the value of $U_{\text{eff}}$, de facto an empirical parameter. Attempts to derive the $U$ value in more rigorous way have been reported in the literature [67]. For the case of the [AlO$_4$]$^0$ defect in silica, it has been shown that the method gives the correct solution, although one has to use a specific value of $U$ which is thus fitted to reproduce existing experimental data [59].

A very similar case is that of a self-trapped hole in SiO$_2$ [68, 69]. This defect consists of a hole trapped at a 2$p$ non-bonding orbital of a single O atom bridging two Si atoms, and is the result of the excitation of a valence band electron to the conduction band, Fig. 1.1. A nearly quantitative agreement between measured and computed hyperfine coupling constants and $g$-factors is found only with a hybrid functional where 50% of the Hartree-Fock exchange is mixed in with the DFT exchange [68]. Standard hybrid functionals, like B3LYP, or pure DFT functionals, like BLYP, fail completely in describing the defect as they provide a delocalised picture, Fig. 1.1.

There is a general problem in the description of these systems since one does not know a priori the amount of Fock exchange that one has to introduce in the hybrid functional, nor can a universal functional describe all systems at the same level of accuracy. Thus, the predicting power of theory for the identification of the spin properties of defects in insulators and semiconductors is presently rather low, and careful comparison with experiments, when exist, is highly desirable.

The example of an Al impurity in SiO$_2$ and of the self-trapped hole, and several similar cases reported in the literature [70, 71], open a more general question for the use of DFT in the description of the oxides and their defects: is the treatment of impurity atoms, missing atoms, oxygen exchange (vacancies formation), etc. sufficiently reliable? In terms of energetics, can one reach the so called “chemical accuracy” (typically an error of 1 kcal/mol in reaction energies) [9]? To what extent errors in the description of the position of the defect states, in the localisation of holes and electrons, reflect in the energetics of the surface reactions? Of course the answer to these questions is not universal, and depends critically on the system considered. In the following, we discuss a particular kind of defect, the oxygen vacancy, in very different oxides and we comment on the results of various theoretical approaches in describing both the electronic structure and the energy changes associated with the formation of such defects.
1.4 The Oxygen Vacancy in Insulating and Semiconducting Oxides

In this section, we will focus on the nature of oxygen vacancies ($V_O$) in six oxides which are representative of insulating and semiconducting oxides. In particular we will consider three wide-gap oxides, SiO$_2$, MgO and ZrO$_2$, and three typical semiconducting oxides, TiO$_2$, NiO, and WO$_3$. The corresponding band gaps are 8.0–10.4 eV (SiO$_2$), 7.8 eV (MgO), 5.0 eV (ZrO$_2$), 3.0–3.3 eV (TiO$_2$), 3.7–4.0 eV (NiO), 2.7–2.8 eV (WO$_3$) [72]. Oxides are also often classified also as “non-reducible” or “reducible”, depending on the facility to remove oxygen from the structure and to form oxygen-deficient structures or even non-stoichiometric compounds [73]. In general, transition metal cations change their oxidation state more easily than main-group elements and belong to the second category. However, we will see that there are transition metal (TM) oxides like ZrO$_2$ where the cost to remove oxygen is not much different from that of other oxides of the main group elements. The discussion is based mostly on data obtained for bulk vacancies, but the general concepts apply as well to vacancies located on the oxides surfaces.

Fig. 1.1 Geometrical structure and spin distribution in a cluster model of a self-trapping-hole center in SiO$_2$. a UHF; b B50-LYP; c B3LYP; d BLYP. Blue Si; red O; yellow H. In grey is shown the spin density.
A source of uncertainty in the comparison of the formation energy, $E_f(VO)$, in various oxides comes from the reference used. $E_f(VO)$ can be computed with respect to the free O atom or to half the energy of the $O_2$ molecule. This latter quantity varies considerably from method to method. In fact, the dissociation energy of gas-phase $O_2$ is 5.23, 6.08, and 6.10 eV with B3LYP, PW91 and PBE functionals, respectively (using the 6-311+G* basis set); the experimental value is 5.18 eV [74]. Thus, while the hybrid B3LYP functional properly describes the strength of the O-O bond, other DFT methods tend to overestimate it. Of course, the problem is less severe if the processes under investigation are surface chemical reactions where oxygen is extracted from the surface and incorporated in a molecular compound. If the bonding of oxygen to the molecular species is also over-estimated by about the same amount then the two errors cancel. Still, this aspect should be considered when comparing computed reaction enthalpies at various DFT levels. It certainly has to be taken into account in the evaluation of $E_f(VO)$. In this chapter we will make reference to the energy required to remove an isolated oxygen atom in its $3P$ ground state, using the values reported above to correct values reported with respect to $\frac{1}{2} O_2$.

In general, an oxygen vacancy in an oxide material can have a magnetic ground state and the assumption that the ground state of this defect is diamagnetic has no foundation. Indeed, in many cases there is a clear preference for forming localised electrons and high spin states. Thus, it is important in the calculation of $V_O$ centers to check both magnetic and nonmagnetic solutions by performing spin-polarised calculations. Notice that this is true also for surface defects and that here the point is crucial in determining the reactivity of the defective surface. In fact, the presence of unpaired electrons trapped at specific defect sites confer a special reactivity to the surface and activates radical mechanisms in bond breaking that are otherwise not possible.

### 1.4.1 Non-reducible Oxides: SiO$_2$, MgO, ZrO$_2$

Silicon dioxide, magnesium oxide, and zirconium dioxide are three oxide materials with broad applications in a variety of technologies. They have in common a high band gap (>5 eV) and can be classified as insulating materials. The band gap decreases considerably going from SiO$_2$ (8–10 eV) to MgO (7.8 eV) and ZrO$_2$ (5.0 eV) [72]. This can in principle induce changes in the nature of the defective material and also in the cost of removing an oxygen atom from the structure.

The electronic structure of $V_O$ in crystalline or amorphous SiO$_2$, is a direct consequence of the covalent polar nature of the Si-O bond. SiO$_2$ in fact forms, in its glassy form, an infinite continuous network resulting from the linkage via bridging oxygens of SiO$_4$ tetrahedra. The high covalent nature of this oxide is shown, for instance, by the directionality of the bond around the Si and the O atoms. Si, tetra-valent, forms tetrahedral structures while O, divalent, form two bonds with an angle of about $140^\circ$, typical of covalent O-containing molecules (e.g. water).
Therefore, in SiO$_2$ the removal of an O atom from a $\equiv$Si-O-Si$\equiv$ linkage results in two Si dangling bonds, $\equiv$Si$, which recombine to form a $\equiv$Si-Si$\equiv$ covalent bond with two electrons occupying a localized state with Si-Si bonding character [75–77]. The process is accompanied by a strong geometrical relaxation and the Si-Si distance decreases from 3.06 Å in the regular lattice to 2.3–2.5 Å in the defect. The associated electronic excitation involves localized excitations from a $\sigma$ bonding to a $\sigma^*$ anti-bonding state that forms in the gap of the material [75]. The large structural relaxation is also the source of relatively large discrepancies in the estimation of $E_f(\text{VO})$ in $\alpha$-quartz. In fact, different models (bare or embedded clusters, supercells, etc.) provide different values of the relaxation, and this can be quite important in the determination of $E_f$. Even larger oscillations are found in the formation energy of an oxygen vacancy in glassy silica, due to the statistical distribution of bond lengths and bond angles present in the structure.

Table 1.1 reports some of the computed $E_f(\text{VO})$ values for crystalline SiO$_2$ [78–81]. All these values have been obtained for bulk silica; due to the continuous network structure of the material, very similar formation energies are also expected for the surface. However, it should be mentioned that Monte Carlo and DFT calculations have shown that, unlike bulk counterparts, the Si-Si configuration of surface oxygen vacancies is likely to be unstable due to the high tensile strain induced [82]. At the highest level of theory used, $E_f(\text{VO})$ in SiO$_2$ is around 9 eV. Calculations performed on various conformations of a $\text{VO}_x$ center in amorphous

<table>
<thead>
<tr>
<th>Oxide</th>
<th>Model</th>
<th>Method</th>
<th>$E_f$ (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MgO(100) surface</td>
<td>Embedded cluster</td>
<td>DFT (B3LYP)</td>
<td>9.1 [97]</td>
</tr>
<tr>
<td></td>
<td>Periodic supercell</td>
<td>DFT (PW91)</td>
<td>9.1 [99]</td>
</tr>
<tr>
<td></td>
<td>Embedded cluster</td>
<td>DFT (PW91)</td>
<td>9.3 [99]</td>
</tr>
<tr>
<td>MgO bulk</td>
<td>Periodic supercell</td>
<td>DFT (PW91)</td>
<td>10.1 [85]</td>
</tr>
<tr>
<td>SiO$_2$ bulk</td>
<td>Periodic supercell</td>
<td>DFT (PW91)</td>
<td>9.6 [79]</td>
</tr>
<tr>
<td></td>
<td>Embedded cluster</td>
<td>DFT (B3LYP)</td>
<td>7.6 [86]</td>
</tr>
<tr>
<td>ZrO$_2$ bulk</td>
<td>Periodic supercell</td>
<td>DFT (B3LYP)</td>
<td>9.8 [87]</td>
</tr>
<tr>
<td></td>
<td>Embedded cluster</td>
<td>DFT (B3LYP)</td>
<td>8.7 [88]</td>
</tr>
<tr>
<td>TiO$_2$(110) surface</td>
<td>Embedded cluster</td>
<td>DFT (B3LYP)</td>
<td>5.3 [89]</td>
</tr>
<tr>
<td></td>
<td>Periodic supercell</td>
<td>DFT (PW91)</td>
<td>6.0 [133]</td>
</tr>
<tr>
<td></td>
<td>Periodic supercell</td>
<td>DFT (PW91)</td>
<td>6.2 [134]</td>
</tr>
<tr>
<td></td>
<td>Periodic supercell</td>
<td>DFT (PBE)</td>
<td>6.5 [135]</td>
</tr>
<tr>
<td>TiO$_2$ bulk</td>
<td>Periodic supercell</td>
<td>DFT (B3LYP)</td>
<td>7.4 [90]</td>
</tr>
<tr>
<td></td>
<td>Periodic supercell</td>
<td>DFT (PBE)</td>
<td>7.3 [90]</td>
</tr>
<tr>
<td>NiO(100) surface</td>
<td>Periodic supercell</td>
<td>DFT (B3PW)</td>
<td>7.0 [91]</td>
</tr>
<tr>
<td></td>
<td>Periodic supercell</td>
<td>DFT+U (PW91)</td>
<td>7.4 [91]</td>
</tr>
<tr>
<td>WO$_3$ bulk</td>
<td>Periodic supercell</td>
<td>DFT (B3LYP)</td>
<td>7.7 [92]</td>
</tr>
<tr>
<td></td>
<td>periodic supercell</td>
<td>DFT (PBE)</td>
<td>7.0 [93]</td>
</tr>
</tbody>
</table>

$E_f$ is computed with respect to the removal of a neutral O$^-$ P atom.
silica show changes in $E_f$ of up to 2 eV related to the different Si-Si distances [77]. This holds true also for the charged variant, the so called $E'$ center (or $V_{O}^{+}$), a defect which undergoes a strong geometrical relaxation, with a single electron localised on a three-coordinated Si atom, $≡Si$ [83]. The spin localisation is proven by theory as well as by EPR experiments [84].

MgO is a prototypical ionic material. It has a cubic NaCl-like structure, where the cations and the anions are in an octahedral environment. Thus, no bond directionality is present, and the charge separation is the consequence of a high Madelung field. This results in a completely different electronic structure of the $V_O$ center compared to SiO$_2$. In MgO a missing O atom from the bulk or from the surface results in two trapped electrons localized in the cavity [94]. The driving force for the localisation of the electrons is the Madelung potential of the highly ionic crystal. The place occupied by the $O^{2−}$ anion in the regular lattice is taken by two “free” electrons in the defective crystal, and the energetic cost of the vacancy formation is minimised. Still, this cost is rather large. $E_f(V_O)$ on the MgO(100) surface is of the order of 9 eV, Table 1.1 [95–99]. $E_f(V_O)$ on MgO is not strongly method dependent. The formation energy decreases considerably when one goes from the regular sites of the MgO(100) terrace to low-coordinated step, edge, corner and kink sites [95]. Here the reduced Madelung potential results in the easier removal of the O atom.

The formation of a $V_O$ center on the MgO surface results in a very small local relaxation (the crystalline potential is only moderately perturbed and the distances around the vacancy change by less than 1 %) [100]. This is completely different from the case of SiO$_2$ and, again, reflects the different nature of the bond in the two materials.

In MgO the electrons trapped at $V_O$ give rise to typical excitations [101]. When the centers are located at the surface of the oxide the transitions fall in the visible region of the spectrum, thus changing the colour of the material. This is why these centers are often called F centers, from the German word for color, Farbe. These centers can also be generated in a charged state, formally by removing an $O^{−}$ ion, with formation of a paramagnetic $F^+$ center (or $V_{O}^{+}$). The generation of F or $F^+$ centers results in new states in the band gap of the material (the gap is about 6.7 eV for the surface [102]). Recently, accurate scanning tunneling spectroscopy measurements on oxygen vacancies created by electron bombardment of MgO thin films have proven the existence of the localised occupied and empty states associated with the missing oxygen in the band gap of the material [103]. See also Chap. 2. Comparison with DFT calculations has allowed an assignment of the observed features to F and $F^+$ centers generated at low-coordinated sites where the formation energy of the vacancy is lower. Furthermore, using non-contact AFM measurements it has been possible to show that a surface $V_O$ center has a strong attractive potential for metal atoms and clusters and acts therefore as a preferential nucleation site, Fig. 1.2 [104]. See also Chap. 7 for a more detailed description of the experimental position.

Basically all theoretical methods, from HF to LDA, give the same correct qualitative picture of electrons trapped at the MgO oxygen vacancy with the
appearance of new states in the gap. The energetic gain associated with localisation of the electron pair in the cavity is sufficiently high that different levels of theory give similar descriptions of the defect. The overestimation of the band gap in HF or its underestimation in LDA (which reflects a more ionic nature of the oxide in HF compared to LDA) does not affect the electronic nature of the defect. This holds true also for the positively charged variant of the oxygen vacancy, the F+ center [100]. This paramagnetic defect has recently been observed on MgO thin films and characterized by EPR experimentally [105] and theoretically [106]. The electronic structure and the spin distribution of this particular center do not seem to depend critically on the method used.

Thus, from the results reported on SiO2 and MgO one can conclude that the electronic structure of defects in wide gap insulators is not particularly affected by the choice of the exchange-correlation functional used, despite the fact that these provide quite different values of the band gap.

What is most affected by the choice of the functional used is the relative position of valence and conduction band states. In a similar way, the position of defect states in the gap may depend on the method. If a chemical reaction involves one of these states (as often is the case) than the corresponding energy changes will also be affected by the position of these impurity levels. A state higher in the gap will exhibit more “basic” or donor character, and vice versa.

We consider now the case of ZrO2. Zirconium dioxide is currently used, for example, as a solid electrolyte in oxygen sensors [107] and solid oxide fuel cells operating at low temperatures [108], in thermal barrier coating applications [109], nuclear waste confinement [110], as a gate dielectric material in metal-oxide semiconductor devices (generally in combination with hafnia) [111], and as a

---

Fig. 1.2 Dependence of frequency shift and interaction energy, respectively, on tip-sample distance. a Shift of the resonance frequency of a Pt0.9Ir0.1 tip on a regular MgO surface (squares) and above a neutral V0 center (F0 defect site) (crosses). Experimental data are derived from constant height measurements. The frequency shift is a direct consequence resulting from potential gradients between tip and sample. The integration of the frequency shift is related to the potential energy. b Interaction energy of a Pt4 cluster above the O site of an MgO surface (rectangles) and above an F0 defect center (crosses) calculated by DFT. c The spillover of the electron charge density of an F0 center calculated by DFT. d The Pt4 cluster above the MgO surface (blue Mg, red O, grey Pt). Adapted from [104]
catalytic support medium [112]. The room temperature stabilisation of the high symmetry polymorphs of ZrO$_2$ is commonly achieved either by transition metal doping or by preparation of nanocrystalline phases. With the B3LYP functional the band-gap of ZrO$_2$ at the $\Gamma$ point is approximately 5.8 eV. This is higher than the measured bandgap of 4.2 eV using electron energy loss spectroscopy (EELS) [113] but within the range measured using vacuum-ultraviolet (VUV) absorption spectroscopy (5.0–6.6 eV) [72, 114].

The removal of an O atom from the bulk of ZrO$_2$ creates a doubly-occupied defect energy level within the band gap, approximately 3.3 eV above the VB. Fig. 1.3 [87]. Figure 1.3 also shows the DOS of the system around the Fermi level and, in the inset, an isosurface of the electron density projected onto this defect state [87]. The state is highly localised at the vacancy site with some density on two zirconium atoms around it. In this respect, it closely resembles a neutral V$_O$ center in alkaline-earth oxides, and has the typical structure of an F center [44]. Structural rearrangements are restricted to the atoms immediately surrounding the vacancy site, in particular the nearest neighbor Zr atoms move radially from the vacancy site. This is reminiscent of the V$_O$ structure in MgO.

After geometry relaxations $E_f$ for this defect is 9.6 eV, very similar to that of a neutral F center in MgO [115] and much larger than that of the same defect in reducible oxides like TiO$_2$ or CeO$_2$ [116] (see below). In this respect ZrO$_2$ must be classified as a non-reducible oxide, like SiO$_2$ and MgO. The high formation energy of an O vacancy suggests that the number of these defects in thermodynamic equilibrium should not be particularly high. Of course, here we are defining as a reducible oxide a material that easily loses oxygen with consequent change of

![Figure 1.3](image)

**Fig. 1.3** Band structure (left) and density of states (right) for ZrO$_2$ with a neutral oxygen vacancy (V$_O$). *Inset* An iso-level of the electron density projected onto the defect state (red O; grey Zr). The zero of energy has been set to the highest occupied level. Adapted from [87]
oxidation state of the metal cation. ZrO$_2$ is not easy to reduce by this mechanism, while it can be reduced by addition of reducing agents like hydrogen [117].

Removing one electron from neutral VO leads to the VO$^+$ paramagnetic defect, Fig. 1.4. The corresponding electronic state splits into two spin-separated components, a singly-occupied spin-up state at about 3.3 eV above the VB and a spin-down unoccupied component just below the CB [87]. Spin density plots show quite unambiguously that the unpaired electron is located in the vacancy, as for F$^+$ centers in alkaline-earth oxides [44]. This is another sign that the Zr$^{4+}$ ion has little tendency to change its oxidation state to Zr$^{3+}$. The geometry relaxations are significantly different to the neutral case, with the Zr atoms moving away from the vacancy center leading to a strong asymmetric site where the electron is trapped.

The fact that the excess electrons present when an oxygen vacancy is created are localised in the cavity and that the corresponding doubly occupied energy levels lie high in the gap is found also when standard GGA functionals are used, and confirms the fact that the nature of the O vacancy in wide gap oxides is not strongly dependent on the computational method used. For an extended analysis of the nature of V$_O$ in the bulk and on the surface of ZrO$_2$ the reader is referred to the extensive review of Ganduglia-Pirovano et al. [58].

1.4.2 Reducible Oxides: TiO$_2$, NiO and WO$_3$

Among the class of reducible oxides TiO$_2$ is most likely the material that has attracted most interest. Titanium dioxide, either in the form or anatase or rutile, is
used in catalysis as a support of metal particles (for instance, nanosized gold particles supported on titania are active catalysts) [118–120]. But the largest interest in this material is related to its photo-catalytic properties and in general its capability to harvest solar light and transform it into a source of chemical energy. The reactive properties of titania are related to the nature and concentration of defects present in the bulk. These, in turn, determine the color of the sample and its conducting properties. The level of defectivity present in TiO₂ depends on many factors, including sample preparation and sample history [121]. This oxide is representative of wide gap semiconducting oxides, having a band gap of 3–3.2 eV [72, 122]. TiO₂ has a degree of ionicity intermediate between fully ionic (Ti⁴⁺ and O²⁻) and fully covalent; a rough estimate indicates a charge of about +2 on Ti and –1 on O. Compared to MgO and SiO₂, its properties are dominated by the presence of 3d orbitals on Ti at relatively low energy. In the stoichiometric form, TiO₂ is a d⁰ oxide and the 3d levels are unoccupied. When an O atom is removed, the presence of the 3d levels opens other possibilities to redistribute the electrons involved in the bonding with respect to non-reducible oxides. This is why the creation of an oxygen vacancy in TiO₂ leads to structures and properties that differ substantially from those described above for MgO, SiO₂ and ZrO₂.

Experiments show that O vacancies in TiO₂ introduce defect states in the band-gap about 0.8 eV below the conduction band [123]. These states can be assigned to Ti³⁺ 3d⁰ ions based on EPR and XPS-UPS experiments [124–126]. Thus, the electrons associated with the removed O atom are neither trapped in the cavity by the strong electrostatic potential (as in MgO or ZrO₂) nor do they lead to the formation of a direct Ti-Ti bond (as found in SiO₂). Since the 3d orbitals are rather localized, the addition of one electron to the Ti⁺⁺ ion corresponds to a formal change of oxidation state, from +4 to +3. Exactly the same situation can be obtained by doping with heteroatoms, such as the addition of alkali atoms [127] or hydrogen [130]. These species form M⁺ cations or H⁺ protons while the valence electron is trapped at a Ti³⁺ ion. The formation of Ti³⁺ species results in three observable consequences: a shift in the core level binding energies of the reduced Ti atoms [128], the formation of paramagnetic centers (detectable by EPR), and the presence of new states at about 0.8 eV below the bottom of the conduction band [123]. These arguments apply as well to bulk as to surface vacancies.

The theoretical description of the Ti³⁺ states however is a delicate issue and depends critically on the method used. Standard DFT provides a Kohn-Sham band gap for TiO₂ that is much smaller than the experimental one, about 2–2.5 eV [129, 130]. The Ti³⁺ states are located at the bottom of the CB, Fig. 1.5, and not in the gap, at variance with the experiment. Not surprisingly, these states with CB character are completely delocalised over several Ti ions. Localisation on one or a few Ti³⁺ ions is obtained when self-interaction corrected functionals are used (such as hybrid functionals or the DFT+U method).

The ground state of VO is magnetic, with two unpaired electrons occupying two different Ti ions [131]. The spin localization is connected to an important structural deformation (polaronic distortion) which is obtained only using self-interaction corrected functionals [130]. The nature of Ti³⁺ states, partly or fully localized,
however, is still matter of debate, and different interpretations are provided based on different experiments. Theory in this respect is in a difficult position, as the exact nature of the VO defect center depends critically on the amount of exact exchange included in a hybrid functional or on the value of the U parameter used in DFT+U, Fig. 1.6 [90]. The situation is complicated by the fact that different solutions, with the unpaired electrons fully localised on a single Ti ion or on a few Ti ions, are very close in energy, often separated by a few tens of meV [90]. In this respect, temperature effects are also very important as experiments performed at different temperatures can monitor rather different situations where the unpaired electrons are hopping between different sites. Experimental issues associated with TiO$_2$ surfaces can be found in Chaps. 3, 4, 8, 14. Moreover, defects on TiO$_2$ powders are discussed in Chap. 9, EPR of TiO$_2$ in Chap. 10, with TiO$_2$ use in resistive switching being discussed in Chap. 13.

It is not surprising that also the formation energy, $E_f(V_O)$, in TiO$_2$ is affected by the computational method used. As for the SiO$_2$ case, a direct comparison is not so simple as $E_f$ (and in general the adsorption properties) depend on several details of the calculation, defect density, number of slabs, pseudopotential used, energy cutoff, amount of relaxation included, etc. [132]. One reference calculation performed at the B3LYP level using cluster models of the rutile surface reports an estimated $E_f$ of about 5.3 eV [89]. Periodic supercell calculations give values on the order of 6–6.5 eV [133–135]. This shows that removing an oxygen atom from bulk rutile has a lower cost than in SiO$_2$ or MgO, consistent with the reducible character of this oxide. Related modelling of steps on oxide surfaces is described in Chap. 6.

NiO is an antiferromagnetic oxide with a cubic structure and a band gap between 3.7 and 4.3 eV [72, 136–139]. Partially filled, narrow 3$d$ bands are present above
the O 2p band. Exchange of electrons from these states forms the basis of the special and rich chemistry of this oxide, but the description of narrow band transition metal and rare earth oxides with standard DFT methods presents severe problems. Standard LDA or GGA functionals give, for instance, a completely wrong electronic structure for NiO. A pragmatic way to circumvent the problem is to use hybrid functionals (e.g. B3LYP or HSE) or other self-interaction corrected functionals [140, 141]. This leads to much better estimates of the band gap for narrow band materials [56, 58, 142]. The other approach which is often used is DFT +U [143]. With the B3PW functional the band gap ofNiO is 3.93 eV [91].

The neutral oxygen vacancy on the surface of NiO has been studied with hybrid functionals. The removal of an O atom from the NiO(100) surface is accompanied by the appearance of new states in the gap of the material, as shown in Fig. 1.7 [91]. The two electrons left by the removed O are partially distributed inside the cavity, but to a larger extent are localised over the Ni atoms around the vacancy. This is shown also by the charge density plots of the defect states, Fig. 1.8. The situation is thus intermediate between that of MgO (full localization in the vacancy) and of TiO2 (reduction of the oxidation state of the metal cations). This picture is found both using hybrid functionals and the DFT+U approach. The formation energy of VO at the NiO(001) surface is about 7.0–7.5 eV, depending on the
method used [91]. These values are about 2 eV smaller than on MgO which is consistent with the reducible nature of NiO compared to MgO.

The last example considered is that of WO$_3$, a key material in several important applications, from smart windows technology to dye-sensitised solar cells, from sensors to photo-electrochemical water splitting and heterogeneous catalysis [144–146]. WO$_3$ is composed of corner sharing regular octahedra and at room temperature assumes a monoclinic structure stable from 17 to 330 °C, Fig. 1.9 [147]. Experimentally, the band gap of WO$_3$ from optical, photocurrent and photoemission measurements varies from 2.5 to 3.2 eV, but most of the results are from 2.6 to

**Fig. 1.7** a Projected DOS curves and b band structure for an O vacancy on the surface of NiO (100). B3PW results on a $2 \times 2$ supercell. In b the band structure for majority and minority spin components are shown. Adapted from [91]

**Fig. 1.8** Charge density plots of the defect states for an O vacancy at the NiO(100) surface (B3PW results, $2 \times 2$ supercell); the left plot is in the surface plane; the right one in a plane normal to the surface through the vacancy and three neighbouring Ni ions. The separation between consecutive iso-density curves is 0.002 a.u. Adapted from [91]
3.0 eV [72, 148]. B3LYP gives a direct Kohn-Sham band gap of 3.13 eV, Fig. 1.9, only slightly overestimated compared to the experimental values [149].

The oxygen vacancy in WO$_3$ is the center responsible for the electrochromic behaviour of this material. Fully oxidised crystalline films do not show any colouration while the colouration becomes intense with increasing oxygen deficiency. Colouration is accompanied by a change in the electrical conductivity from semiconducting to metallic behavior, reinforcing the idea of the involvement of V$_O$ centers. It has been proposed that an optically or electrically induced change of the charge state of V$_O$ should result in a corresponding change in absorption properties and coloration. The origin of electrochromism is thus related to the nature of V$_O$ in WO$_3$, and this center has been studied theoretically with DFT methods [150, 151].

Given the monoclinic crystal structure of WO$_3$, a bulk oxygen vacancy can be created in an -O-W-O-W-O- chain along the x, y, or z crystallographic axes, Fig. 1.10. This gives rise to rather different electronic structures [92, 152]. Oxygen removal results in two excess electrons, which can be either paired up in a singlet (closed shell) spin state or unpaired in a triplet spin state. In the diluted case (2 % doping), when V$_O$ in the -O-W-O-W-O- chain is along the x direction (see Fig. 1.10), the singlet is slightly more stable than the triplet (0.2 eV) [92]. The charge density is mainly localised at the vacancy site, with small tails of the wave function on the 5d orbitals of the undercoordinated W atoms of the cavity. The defect can be schematically described as W$^{6+}$/V$_O$(2e$^-$)/W$^{6+}$ and, in this respect, is reminiscent of the electron localisation found in wide gap oxides like MgO or ZrO$_2$. The associated electronic state is well separated from the bottom of the CB, forming a moderately flat band in the gap at about 0.5–1.0 eV below the CB minimum. At low levels of defectivity (2 %), WO$_{3-x}$ still has semiconducting character. A higher V$_O$ concentration (4 %) has a dramatic effect on the electronic structure of WO$_{3-x}$ as the defect states become rather dispersed and merge with the conduction band, leading to metallic character. When the vacancy is created along the y direction the V$_O$ ground state is triplet, 0.15 eV more stable than the lowest singlet state. Two defect bands, occupied by one excess electron each, are present in the band

![Fig. 1.9 Geometric structure (left) and band structure (right) of RT monoclinic WO$_3$. Red O; green W. Adapted from [149]](image)
gap. One of these bands, more dispersed, merges with the CB leading to a metallic character. In contrast to a VO defect created along the x direction, the spin density on the two neighboring W atoms is very asymmetric (1.40 vs. 0.23 e\(^{-}\)), indicating the formation of W\(^{4+}\)/W\(^{6+}\) pairs (schematically W\(^{4+}\)/VO(0e\(^{-}\))/W\(^{6+}\)). The last case is that of a VO defect generated along the perpendicular z direction. Here the triplet state is 0.30 eV more stable than the singlet, again showing a magnetic nature of the defect. However, in contrast to both the x and y directions, both defect bands are flat and the state is localised. The spin density is symmetrically distributed on the two undercoordinated W atoms (W\(^{5+}\)/VO(0e\(^{-}\))/W\(^{5+}\))\[92\].

To summarise, there is a clear dependence of the properties of O vacancies on the crystallographic direction of the -W-O-W- chain where the atom is removed. On the other hand, the energy cost to create a vacancy is less dependent on the crystallographic direction. B3LYP calculations on a 32 atoms supercell give formation energies of about 7.7 eV; PBE calculations on the same supercell size give values between 6.9 and 7.1 eV, Table 1.1\[92\]. This suggests that upon thermal treatment a statistical distribution of all types of vacancies will form. Notice that the formation energy of the vacancy in bulk WO\(_3\) is comparable to that of other reducible oxides like NiO and TiO\(_2\), Table 1.1.

### 1.5 Conclusions

The nature of an oxygen vacancy in the bulk or on the surface of an oxide can be considered as a fingerprint of the oxide electronic structure. This depends on the redistribution of the two electrons associated with the lattice O\(^{2-}\) ion once the O atom is removed from the structure. In ionic oxides like MgO the strong Madelung
potential helps in confining the extra charge in the cavity formed by removing the O atom. The electron (or electrons) do not reduce the neighboring cations, and remain localised in the centre of the void. Little distortion occurs because the electrostatic interaction is basically preserved and no new direct bond is formed. In mixed covalent-ionic oxides, like SiO$_2$, where the Madelung field is smaller because the atomic charges are smaller, the two electrons localise on hybrid orbitals of the two Si neighbors forming two dangling bonds that point towards the cavity formed by the missing oxygen. This leads to the formation of a direct Si-Si bond with consequent distortion of the region around the defect. The third situation is that of reducible oxides where the extra electrons are neither localised in the cavity nor used to form a new localised bonding electron pair. Here the two electrons are transferred to the low-lying empty d or f states of the metal cations, with formal reduction of their oxidation state. Since the two electrons can be transferred to two different cations, this results in paramagnetic defects and a magnetic impurity.

In all cases, the creation of the defect results in new electronic states in the band gap of the oxide. Depending on its nature, reducible or non reducible, these states are deeper in the gap or closer to the conduction band. In particular, in non reducible oxides like MgO, SiO$_2$, and ZrO$_2$ the filled defect states are near the center of the gap; in reducible oxides like TiO$_2$ the defect state is closer to the conduction band. Of course this is not a rule, as the position of the defect state also depends on other factors like the external electrostatic potential, the level of localisation etc.

These arguments also apply to surface vacancies. On the surface, the presence of these defects can alter completely the reactivity of the material. In fact, the presence of extra electrons less strongly bound than the electrons of the O 2$p$ valence band results in a much stronger ability to activate chemical bonds and induce chemical reactions. In this respect, a very critical parameter is the number of these defect centers in the native material. This, in turns, depends on the formation energy of the defect.

The microscopic understanding of the surface reactivity requires an appropriate description of the electronic structure and the properties of these defect centers. Here things are more delicate, as the exact nature of oxygen vacancies is not easy to describe at both experimental and computational levels. Experiments have to face the problem of using sophisticated spectroscopies or microscopies which are sensitive to a very small number of centers; in this respect, great advances have become possible thanks to the use of local probes like STM and AFM that often provide atomic resolution images of the defects. Also local spectroscopies like STS have considerably improved our understanding of the electronic structure of specific surface defects. In this scenario, a non-negligible role is that provided by electronic structure calculations.

With the advent of density functional theory, the treatment with first principles approaches of systems that have practical interest has become possible. This includes defects and oxygen vacancies on the surface and in the bulk of crystalline and amorphous materials. Despite the intrinsic complexity of these systems, DFT
calculations represent a formidable tool to interpret experiments and provide a solid basis for the assignment of a given spectral feature to a specific defect. DFT is also essential to model surface reactions and the mechanisms of heterogeneous catalysis. This often involves the presence of oxygen vacancies or the formation of oxygen vacancies during the reaction. The accurate description of these processes is, however, still a challenge. While the general trends are usually solid and can be trusted, the specific values of a given formation energy strongly depend on the form of exchange-correlation functional used. This holds true not only for the energy aspects but, sometimes, also for some details of the electronic structure (position of the defect states in the gap, level of localization, formation of paramagnetic centers, etc.). Substantial activity is dedicated to solve these problems, for instance by going to higher levels of theory like many-body approaches (GW) [153, 154] or full configuration interaction quantum Monte Carlo [155]. The use of these approaches is presently limited by their high computational cost but there is little doubt that in the future these methods will provide a way to obtain more accurate descriptions of solids and their defects. The role of electronic structure theory in the design of materials and their surfaces is also expected to grow.

Acknowledgments I am indebted to all coworkers and collaborators that over the years contributed to this work: Prof. Thomas Bredow, Prof. Cristiana Di Valentin, Dr. Anna Maria Ferrari, Prof. Hans-Joachim Freund, Prof. Elio Giamello, Dr. Livia Giordano, Prof. F. Illas, Prof. Anna-bella Selloni. Financial support from the Italian MIUR through the FIRB Project RBAP115AYN “Oxides at the nanoscale: multifunctionality and applications” and the COST Action CM1104 “Reducible oxide chemistry, structure and functions” are gratefully acknowledged.

References

36. M. Born, Z. Physik 1, 45 (1920)
Defects at Oxide Surfaces
Jupille, J.; Thornton, G. (Eds.)
2015, XVI, 462 p. 202 illus., 95 illus. in color., Hardcover
ISBN: 978-3-319-14366-8