Topical Review

The origin of two-dimensional electron gases at oxide interfaces: insights from theory

N C Bristowe\textsuperscript{1,2}, Philippe Ghosez\textsuperscript{1}, P B Littlewood\textsuperscript{2,3,4} and Emilio Artacho\textsuperscript{2,5,6}

\textsuperscript{1} Theoretical Materials Physics, University of Liège, B-4000 Sart-Tilman, Belgium
\textsuperscript{2} Theory of Condensed Matter, Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, UK
\textsuperscript{3} Physical Sciences and Engineering, Argonne National Laboratory, Argonne, IL 60439, USA
\textsuperscript{4} James Franck Institute, University of Chicago, 929 E 57th Street, Chicago IL 60637, USA
\textsuperscript{5} CIC Nanogune, and DIPC, Tolosa Hiriideia 76, E-20018 San Sebastian, Spain
\textsuperscript{6} Basque Foundation for Science Ikerbasque, E-48011 Bilbao, Spain

E-mail: n.bristowe@ulg.ac.be

Received 1 October 2013, revised 4 February 2014
Accepted for publication 4 February 2014
Published 18 March 2014

Abstract

The response of oxide thin films to polar discontinuities at interfaces and surfaces has generated enormous activity due to the variety of interesting effects that it gives rise to. A case in point is the discovery of the electron gas at the interface between LaAlO\textsubscript{3} and SrTiO\textsubscript{3}, which has since been shown to be quasi-two-dimensional, switchable, magnetic and/or superconducting. Despite these findings, the origin of the two-dimensional electron gas is highly debated and several possible mechanisms remain. Here we review the main proposed mechanisms and attempt to model expected effects in a quantitative way with the ambition of better constraining what effects can/cannot explain the observed phenomenology. We do it in the framework of a phenomenological model constructed to provide an understanding of the electronic and/or redox screening of the chemical charge in oxide heterostructures. We also discuss the effect of intermixing, both conserving and not conserving the total stoichiometry.

Keywords: oxide interfaces and surfaces, polar discontinuity, two-dimensional electron gas, lanthanum aluminate, strontium titanate, net charge, charge screening mechanisms

(Some figures may appear in colour only in the online journal)
1. Introduction

Oxide interfaces have received considerable attention within the past couple of decades. From the device perspective, many technological applications depend crucially on interfacial properties, and the fundamental understanding at the microscopic level is often critical for optimizing device performance. In particular oxide interfaces have also offered the opportunity to discover and study completely novel, and often unexpected, physical properties and a wide range of potential device applications have made the LAO–STO interface a highly popular system for study, with several dedicated reviews and perspectives [15–24].

Despite the enormous amount of activity on this system, the origin of a two-dimensional electron gas (2DEG) at the LAO–STO interface remains unclear. Arguably the most popular explanation, called the ‘electronic reconstruction’, is based on the concept of a ‘polar catastrophe’ [25]. The concept uses basic electrostatics (see section 2.7) to show that a net charge at a polar, insulating interface such as LAO–STO produces a finite non-decaying electric field. This diverging electrostatic potential is clearly energetically unstable upon increasing the thickness of the oxide film. In reality the system requires some form of charge compensation. This might occur through an electronic reconstruction (see section 3) whereby the electronic field produces an electronic transfer similar to Zener tunnelling from the surface valence band to the interface conduction band. This process should not be taken too literally. The kinetic process of establishing the 2DEG can be more involved and is beyond the scope of this paper. The polar catastrophe and electronic reconstruction concepts establish the paradigm in an intuitive way. It is interesting, however, that the same mechanism can be seen from a different perspective [26], where it was noted that the isolated neutral interface and surface should have free carriers associated with them to start with, which (partly) annihilate when brought close to each other. It is an alternative view of the same mechanism, with the same predictions for the equilibrium result, but a different perspective on the process.

Alternative explanations suggest that charge compensation may be achieved through chemical modifications or chemical plus electronic (redox) modifications (see section 4). A third category of proposed mechanisms is not based on polarity arguments but instead on extrinsic effects induced by the film growth process such as doping through oxygen vacancies or interface cation intermixing (see section 5). Clearly the origin of the 2DEG is not just of academic interest, but key to understanding what other systems might show similar behaviour and how one might be able to tune device properties.

A large number of experimental and theoretical studies have been devoted to clarifying which mechanism(s) operate in LAO–STO. A recent short review attempted to summarize the results, concluding that the electronic reconstruction mechanism appears to support experiment [21]. The electronic reconstruction mechanism is certainly strongly supported by several experimental observations, most notably the observation of a critical thickness of the LAO film for the onset of interface conductivity [7] which can be tuned [27]. However the electronic reconstruction mechanism does not naturally explain all experimental observations. For example the discovery of a sizable density of (trapped) Ti 3d-like states below the critical thickness [28–32], the lack of surface hole carriers [7] or states near the Fermi level [33, 34], the almost negligible electric field within LAO below the critical thickness [32, 35, 36], and the apparent disappearance of conductivity at any LAO thickness for samples grown at high oxygen partial pressures [37, 38]. It is apparent that the situation is not so clear.

The main aim here is to review the findings for this and related systems in the framework of a simple model; this model describes simply but quantitatively the several effects that contribute to the physics of such systems. In doing so, we hope to better constrain what effects can/cannot explain the observed phenomenology. It is therefore limited in scope and far from any ambition of being exhaustive in the review of the activity around these systems. In section 2 we briefly introduce the concept of polar interfaces, highlighting for the case of LAO–STO the basic theory behind determining the interface net charge and resulting electrostatics. We then begin by exploring the various proposed mechanisms behind the origin of 2DEGs at oxide interfaces. We present a phenomenological model constructed to provide an understanding of the electronic and/or redox screening of the chemical charge in polar heterostructures in sections 3 and 4. In section 5 we discuss the effect of intermixing, both conserving and not conserving the total stoichiometry.

2. Polar interfaces

2.1. The net charge

In the seminal paper of Ohtomo and Hwang [1] the discovery of the two-dimensional electron gas at the interface between
The two nominally insulating oxides, LAO and STO, was rationalized in terms of the electrostatics of the formal charges on each ion, i.e. $-2e$ for O, $+4e$ for Ti and so on, as illustrated in figure 1. From these formal charges, the (001) layers in STO (SrO and TiO$_2$) are charge neutral whilst in LAO (LaO and AIO$_2$) they are $+1e$ and $-1e$ per layer formula unit. At the (001) LAO–STO interface, taking the two unperturbed bulk materials, these layer charges give rise to an imbalance at the interface of $\pm 0.5e/\Theta$. The sign depends on the interface termination, and $\Theta$ is the 2D interface unit cell area corresponding to one formula unit. These basic arguments, built upon the formal ionic charges, are at the heart of almost all subsequent work on this system.

It is tempting to suggest a renormalization of these numbers by introducing covalency factors in an attempt to correct for realistic bonding in these far from ideally ionic ABO$_3$ materials. It is well known that in most perovskites there is a sizable covalent bonding character, especially in the B–O bonds. One would then find after repeating the basic procedure above an arbitrary value of the interface charge smaller than $\pm 0.5e/\Theta$.

The first point that we wish to highlight here is that when determining the net interface charge in these materials it would be incorrect to include covalency considerations such as that just mentioned (see [39] for further discussion). The net interface charge of these band insulators is exactly as obtained from the simple counting of the formal charges on each species irrespective of covalency$^8$. For LAO–STO, the net interface charge is precisely $0.5e/\Theta$ [41, 42].

Let us discuss this further. We outline the arguments proposed in [39–43]. For electrostatics purposes (see the next section) the calculation of the resulting electric field from an interface requires the knowledge of both the area density of free charges and the total polarization, or bound charges (including the electric field response—see the next section) of each material. The subtlety of the problem comes from what is meant by free charges and the definition of the polarization. Remember that we are considering the unperturbed bulk interface, so zero free charges in a strict sense. However the free charge can also be defined as anything unaccounted for by the polarization, and previously called chemical [41] or compositional charge [44]. The key to the problem lies in the definition of the reference unperturbed (zero of) polarization: altering its value will change the free charge by the same amount. In the end it is the sum of the two that matters and is unaffected. Centrosymmetric materials such as LAO and STO allow for a natural definition of zero polarization, the effective polarization then being the change of polarization after some perturbation from the centrosymmetric reference. Taking the polarization of STO and LAO as zero, one can then use a charge counting technique such as the ‘polarization-free’ unit cell method [45] to find a free (or chemical/compositional) interface charge density of $0.5e/\Theta$.

2.2. Polarization and quanta

An alternative viewpoint, that is perhaps more elegant and rigorous, has been proposed [42] which considers the formal polarization instead of the effective value above. The formal polarization is obtained by direct application of the modern theory of polarization [46], for example by taking the ionic cores and charge centre of Wannier functions [47] as point charges. The difference between the two definitions, effective and formal, for the case of LAO–STO is exactly $0.5e/\Theta$: whilst the effective value for LAO and STO is zero, the formal values are $0.5e/\Theta$ and zero, respectively. Within this definition the
free charges are now zero, but the sum of free and bound charges remains $0.5e/\Theta$ and hence the resulting electrostatic analysis is unaltered.

We have neglected to mention so far one more subtlety. In a system of discrete point charges the bulk polarization is defined only up to quanta of polarization. When the definition of the bulk unit cell is displaced along a crystallographic direction and reaches one of the point charges, it disappears there and reappears at the other side of the unit cell. At this point the polarization obtained jumps by a finite value of

$$P_0 = 1 eL/V = 1 e/\Theta,$$

which is called a quantum of polarization, where $L$ is the length of the cell along the direction of the flip, and $V$ is the cell volume (the polarization being the dipole moment per unit volume). Clearly there is no a priori best choice of the unit cell as long as the surface is unknown. The bulk polarization value is therefore only defined modulo polarization quanta. In centrosymmetric perovskites two sets of formal polarization values are allowed,

$$\left\{ \ldots, -2P_0, -P_0, 0, P_0, 2P_0, \ldots \right\},$$

as in STO, and

$$\left\{ \ldots, -3P_0/2, -P_0/2, P_0/2, 3P_0/2, \ldots \right\},$$

as in LAO, since these are the only two sets that are invariant under change of sign. This is the key difference from the effective polarization as defined above, where all centrosymmetric materials were defined to have 0 polarization: now they can have 0 or $P_0/2$, both modulo $P_0$.

2.3. Symmetry and topology

We have approached the problem of determining the polarization from a perspective based on point charges using the centre of charge of the valence band Wannier functions [47]. An alternative approach defines the bulk polarization of an insulator in terms of a Berry phase [46, 49]. This formulation, called the modern theory of polarization, arises from the consideration of the current that flows through the system when adiabatically changing its polarization state. In 1984 it was established by Berry [50] that such an adiabatic change of a Hamiltonian is associated with the appearance of a complex phase in the wavefunction that describes the system as it adiabatically evolves. This phase is of deep physical significance and is called a geometric phase, which, in principle, can take any value, corresponding to an arbitrary value of the polarization. The presence of a symmetry pins the polarization (and corresponding phase) to particular values, and the phase is referred to as topological.

In centrosymmetric insulators like LAO and STO, the two possible values of the polarization 0 and $P_0/2$ (both modulo $P_0$) described in section 2.2 correspond to the two values of the Berry phase, 0 and $\pi$, modulo $2\pi$. It is interesting to make the analogy with other symmetry-protected topological phases (in e.g. topological insulators). Here the relevant symmetry is e.g. time reversal, and the topological phase relates to other properties, e.g. to magnetization or its susceptibility. In the conventional table for topological insulators, one axis corresponds to the relevant symmetry, its entries conventionally being time-reversal and particle–hole symmetries, and their product. Arguably, in this table categorizing the different symmetry-protected topological phases, spatial symmetries should also feature, which would then include the insulators considered in this review. This has already been done for two-dimensional systems [51, 52], referring to the quantization of electric polarization, as discussed here. Interestingly, for other symmetry groups, as for honeycomb planar insulators (e.g. BN), there are more than two sets of consistent values, three in this case [51]. (There are further topological classifications of insulators based on crystal symmetries, but in a different context to that of this review [53].) In analogy with other topological insulators, the interface between two insulators of different topological character in this sense, such as LAO and STO, gives rise to a 2DEG. Of course, such 2DEGs are not topologically protected against any kind of defect breaking the symmetry defining the phase. This is analogous to the situation relating to a $Z_2$ topological insulator: the 2DEG arising on its surface is not protected against defects breaking the time-reversal symmetry, such as magnetic impurities. In this case, the 2DEG is affected by regular disorder and defects.

![Figure 2. Symmetry-protected topological phases. (a) Sketch of a finite Haldane chain, a quantum antiferromagnet of Heisenberg spins with $S = 1$, which displays net $S = 1/2$ at the ends. It is adapted from the figure in [48]. (b) Analogous sketch for a chain of quantum charges, taking delocalized electrons and localized nuclei. The middle panel shows the quantum delocalization of the former, and the corresponding Wannier functions. The net effect is that of half the quantum charge at each end.](image-url)
The analogy with other symmetry-protected topological phases is best illustrated referring to a ‘perspective article’ of Qi in Science in 2012 [48]. Here the concept of a symmetry-protected topological phase is transmitted in terms of the Hal
dane chain, a one-dimensional Heisenberg model of localized quantum spins with \( S = 1 \), interacting antiferromagnetically. In the ground state for a finite chain, the quantum fluctuations of the local spins act in such a way that the chain appears as if without spins, except for a net spin of \( S = 1/2 \) at one end and \( S = -1/2 \) at the other, in spite of their being no \( S = 1/2 \) particles anywhere in the system.

That system is beautifully analogous to a chain of alternating charges, + and − some quantum of charge, which is again analogous to the (100) atomic planes of LAO (see figure 1 of [39]). The net effect is that of half a quantum of charge of either sign at either end, as sketched in figure 2 (the quantum fluctuations in this case manifest themselves in the delocalization illustrated in the figure for the electrons).

2.4. Other interface planes

The discussion above was centred around the heterostructures grown in (001) planes because of most experiments having been performed for those systems, in addition to the simplicity that they offer. The generalization to other growth directions is simple, either using formal charges or using polarization. Using formal charges, the argument goes parallel to the original one of Ohtomo and Hwang [1]. The stacking of [001] atomic layers in the ABO\(_3\) perovskite is AO/BO\(_2\)/AO/BO\(_2\), and thus with formal charges +1, −1, +1, −1, for LAO, while they are all neutral for STO. Similarly the stacking of [111] layers is AO\(_3\)/BO/BO\(_3\)/BO, and therefore with alternating −3 and +3 charges in LAO, and −4 and +4 for STO. For [011] the stacking is ABO/BO\(_2\)/ABO/O\(_2\), and the resulting formal charges are +4 and −4 for both materials. One would thus expect a polarization discontinuity in [001] and [111] but not in [011].

Going back to the polarization discussion, the arguments above generalize to 3D from the 1D set proposed in such a way that for one meaningful value of the polarization there is a whole 3D lattice of \( \vec{P} \) values. In the STO case, and assuming the cubic phase for simplicity, the lattice is simple cubic, with \( P_0 \) as the lattice parameter, and such that \( \vec{P} = 0 \) is a member of the set. For LAO the lattice is also simple cubic, but now the origin \( \vec{P} = 0 \) is at the centre of a cube, and thus not part of the set. This is illustrated in figure 3(a), and it is analogous to the polarization lattices already presented in the pioneering paper of Vanderbilt and King-Smith [43], and to the recent proposal for 2D honeycomb (graphenic) insulators in [51].

In order to predict possible polarization discontinuities of these two materials in interfaces of arbitrary orientation, one selects the direction normal to the interface and projects all \( \vec{P} \) vectors of the lattice for each material onto the particular direction line. This is illustrated in figure 3, using the [001], [011] and [111] directions as examples. The projections proposed in figure 3(a) give rise to the 1D sets for the three directions, shown in figure 3(b), for both the STO (continuum lines and filled circles) and the LAO case (dotted lines and open circles). It is very apparent that both for (001) and for (111) interfaces there are no possible choices of \( \vec{P} \) for both materials such that \( \Delta \vec{P} = 0 \); there is always a discontinuity in the polarization.

Consistently with this discussion and in addition to the well-known 2DEG in the (001) case, a 2DEG has been recently discovered at the (111) LAO–STO interface [54]. In this case, the atomic layers are alternating AO\(_3\) planes and B planes, which for LAO means −3 and +3 layers,
while in STO it means $-4$ and $+4$ layers. There is therefore a discontinuity at the interface corresponding to $0.5e/\Theta_{111}$. Noting that $\Theta_{111} = \sqrt{3}\Theta$ (for the $\Theta$ as defined in section 2.1, for the (001) case) we see that the formal charges argument, giving a chemical charge of $0.5e/(\sqrt{3}\Theta)$, agrees with what was obtained from figure 3, namely, $\Delta P = 0.5 P_0/\sqrt{3}$. While a discontinuity in the polarization exists for (001) and (111) interfaces, this is not the case for the (011) interface for which interfaces with no $P$ discontinuity can be built. This agrees with the consideration from formal charges for which (011) planes of LAO (STO) are made of LaAlO (SrTiO) planes alternating with $O_2$ planes, i.e. alternating planes of $+4e/\Theta$ and $-4e/\Theta$, in both cases. Therefore, as long as an $O_2$ plane always separates ABO planes, there should be no polarization discontinuity.

It is interesting, however, that experiments for (011) interfaces [54, 55] show interface conduction similar to what was found in the (001) case, in spite of the fact that according to the discussion in the previous paragraph there should be no polar discontinuity at the pristine interface, and thus no intrinsic tendency to accumulate free carriers at the interface. The same study [55] arrives at an explanation of the effect by a careful structural study of the interface, which shows that it is not an ideal epitaxial (011) interface, but rather a heavily stepped one, in a jigsaw pattern of (001) steps and terraces. These do present the polar discontinuity discussed above, thereby explaining the appearance of the 2DEG. This raises the next question, as to why this pattern emerges at all if the ideal interface would represent no electrostatic penalty whatsoever. The answer as suggested in [55, 56] is likely to be related to the fact that although the ready-made interface would not be polar, the preformed ideal (011) surface of STO is highly so and requires pre-screening, such as the observed formation of the (001) steps and terraces, before the growth of the film.

Finally, the vector lattice above can also be used to predict the polar discontinuities at stepped or vicinal surfaces, as demonstrated in [57]. The direction normal to the average interface defines the global polar discontinuity, although if the interface is constituted of well-defined steps separated by wide terraces, the contribution of steps and terraces can be distinguished, allowing for the preparation of one-dimensional electron gases along the steps [57].

2.5. Removing the arbitrariness

In as much as we are defining the polar discontinuity at an interface from the bulk polarization values, there is the indeterminacy in the lattice of quanta described above. However, the formal polarization can be argued to give a more defined information for a given surface or interface, even an absolute value, as was already proposed by Vanderbilt and King-Smith [43].

A compelling piece of evidence in this direction is given by the case of the (001) interface between e.g. K$TaO_3$ (KTO) and LaAlO$_3$. Taking the formal charges of $-2e$ for $O$, $+1e$ for $K$ and $+5e$ for Ta, the (001) layers in KTO are $+1e$ (TaO$_2$) and $-1e$ (KO) per layer formula unit. Both LAO and KTO are of the $P_0/2$ kind, and thus one can foresee a non-polar interface. Yet, an ideal epitaxial interface, if considering formal charges, would show two $-1$ layers or two $+1$ layers together at the interface, since the AO planes are $+1$ for LAO, but $-1$ for KTO, and vice versa for the BO$_2$ planes. Such interfaces would give a net chemical charge of $\pm 1e/\Theta$, i.e. $\Delta P = \pm P_0$. In other words, although an interface between KTO and LAO could be made without polar discontinuity, it seems that the ideal epitaxial one is polar. Enhanced carrier densities in interfaces between III–III and I–V perovskites have indeed been predicted in [58].

The fact is that when defining a whole system with explicit surfaces, the absolute polarization is well defined, and can be measured, as the total dipole over the volume of the sample. The surface theorem essentially says that for specific surface terminations (and as long as the system remains insulating) the total polarization obtained from the Wannier charge centres gives the absolute value of the polarization. The absolute value of the polarization is one specific vector from within the lattice discussed in sections 2.2 and 2.4. This is true as long as the periodicity parallel to each surface is respected. This should be understood in a strict sense: if any local source of charge (ion) is added or removed from the surface, it is done periodically, and therefore respecting charge quantization per unit cell. This also holds for interfaces. In the KTO/LAO example, the (001) epitaxial interface, i.e., the one respecting the AO–BO$_2$ plane alternation, gives a polar discontinuity of one whole quantum of polarization. However if we were to grow an interface in which the first BO$_2$ plane of KTO, that is, the TaO$_2$ plane, would grow on a BO$_2$ plane of LAO (the AlO$_2$ plane), there would be no polar discontinuity. That is, the termination defines the point in the polarization set.

2.6. Alternative formulations

Polar surfaces and interfaces have been known for many decades. The description that we have reviewed so far started in a seminal paper [43] and is linked to the modern theory of polarization [46] but there were previous characterizations of such systems. The review of Goniakowski et al [45] gives an excellent account of the varied descriptions of polar surfaces and interfaces. For classical point particles one finds these concepts discussed for related situations in the book by Born and Huang [59]. Tasker, in 1979, offered a characterization and classification of ideally ionic surfaces [60]. The link between Tasker’s characterization and the ideas based on the modern theory of polarization was recently established by Stengel [40].

The treatment for quantum delocalized electrons presents difficulties for some theoretical frameworks, however, related to the problem of how to assign a static charge to the different atoms in a solid. This is the case for instance for the electron counting models of, e.g., Harrison et al [61] by which excess charges were assigned to the cations and anions of heteropolar semiconductors, trying to account for covalency effects. Such techniques were quite successful at a qualitative level when used in conjunction with empirical studies of the electronic structure of materials. The quantitative determination (both from first principles and experimentally) of such quantities however is more problematic, since it involves the assignment of charge to the atoms in the solid, a badly posed problem. It
was attacked with different schemes defining atomic charges, probably the most popular one in recent times being that of Bader [62]. For a brief discussion of the ill-defined nature of static charge schemes in the context of perovskites we refer the reader to [63].

Interestingly, in highly symmetric systems, like cubic LAO and STO, such charge redistribution terms cancel and the final result for the polar discontinuity coincides with the \( P_0/2 \) value established above. For systems in which polarization is not defined by symmetry, it is by no means obvious to us that the calculation of the polarization discontinuity by means of such charge attributions should agree with the results obtained from the calculation of the formal polarization. We will not discuss the methods further here, and we refer the reader to the aforementioned review for a quite comprehensive discussion of such methods [45].

Finally, another take on the matter worth mentioning is that due to Finnis [64] who discusses polar surfaces and interfaces in terms of the theory of thermodynamic excesses, a way of thermodynamically considering relative proportions of the different components in a phase when close to a surface, interface or grain boundary. Such excesses (changes of concentration with respect to the bulk phase) also relate to charge excess, which connects with the chemical or compositional charge discussed in section 2.1.

### 2.7. Electrostatics

Within the macroscopic electrostatics relevant for length scales larger than atomic size, electric fields and charges are related through Gauss’s law, \( \nabla \cdot \vec{D} = \rho \), where \( \vec{D} \) is the electric displacement field, and \( \rho \) is the charge density of free carriers. For an interface it becomes

\[
D^L_z - D^R_z = \sigma, \tag{1}
\]

where \( \sigma \) is the area density of free charges associated with the interface, \( z \) is taken as the direction normal to the interface and \( L \) and \( R \) indicate the materials on either side of the interface. From the definition of the electric displacement field, \( \vec{D} = \varepsilon_0 \vec{\varepsilon} + \vec{P} \) where \( \varepsilon_0 \) is the dielectric permittivity of vacuum, \( \vec{\varepsilon} \) the electric field and \( P \) the polarization, we obtain the change in electric field across the interface, \( \varepsilon_0 (\varepsilon_z^L - \varepsilon_z^R) = \sigma - (P_z^L - P_z^R) \), or

\[
\varepsilon_0 \Delta \varepsilon_z = \sigma - \Delta P_z. \tag{3}
\]

Therefore, as stated in section 2.1, for electrostatic purposes it is only the sum of the interface charge and the polarization discontinuity across the interface that is required to obtain the resultant electric fields. Separating the spontaneous component from the electric field response component of the polarization, we write

\[
\Delta P_z = \Delta P_z^0 + \Delta P_z(\varepsilon), \tag{4}
\]

and, as argued above for LAO–STO,

\[
\sigma - \Delta P_z^0 = P_0/2 = 0.5e/\Theta, \tag{5}
\]

irrespective of whether we use the effective or formal polarization to define the problem, which from now on we denote by \( \sigma \).

The only thing that remains is to define the geometry and electrostatic boundary conditions and to solve for the electric fields self-consistently, including the dielectric responses. The most widely studied LAO–STO geometry is a single thin film of LAO of thickness \( d \) on a TiO\(_2\) terminated STO substrate\(^9\). This amounts to a system of two interfaces with equal and opposite charges, similar to the electrostatics of a capacitor. Assuming open-circuit boundary conditions (\( \varepsilon = 0 \) within the substrate and vacuum), and taking a linear dielectric response for LAO,

\[
P_z(\varepsilon)_{\text{LAO}} = \varepsilon_0 x_{\text{LAO}} \varepsilon_{\text{LAO}}, \tag{6}
\]

we obtain the usual capacitor expression

\[
\varepsilon_{\text{LAO}} = \sigma_{\varepsilon}/\epsilon \tag{7}
\]

where

\[
\epsilon = \epsilon_0 (1 + x_{\text{LAO}}). \tag{8}
\]

In other words, in the absence of electronic or ionic reconstructions, a constant non-decaying electric field exists in LAO due to the polar interface charge \( \sigma_{\varepsilon} \), which is partly screened by the dielectric response of LAO. This dielectric response of LAO in these films has been predicted within first-principles calculations \([65–67]\), and experimentally inferred through surface x-ray diffraction (SXRD) measurements \([68]\) and high angle annular dark-field imaging (HAADF) \([69]\) which revealed an off-centring of the cation positions along \( z \) with respect to the ideal perovskite structure, and through the electrostriction effect \([70]\).

### 3. Electronic reconstruction

#### 3.1. The concept and model

We start with a simple model that describes the ‘electronic reconstruction’ that has been proposed as a mechanism for counteracting the electrostatic build-up that originates when growing a thin film of LaAlO\(_3\) on an SrTiO\(_3\) substrate (LAO on STO) along the [001] direction \([25]\). The model captures the essential physics of the pristine system as confirmed by several first-principles calculations (see for example \([8, 41, 65, 67, 72]\)). It is illustrated in figures 4(a) and (b): the 0.5\( e/\Theta \) of chemical or compositional charge at the interface and surface (corresponding to the half a quantum of formal polarization of LAO; see section 2.1 and \([39]\)) gives rise to an electric field in the film (see section 2.7). This induces a transfer of electrons from the valence band at the surface to the conduction band at the interface (see figure 4(b)). The electron transfer only happens once the potential drop across LAO, \( V = \varepsilon d \), reaches the relevant band gap, \( e\Delta \) (here we will take the gap parameter \( \Delta \) as an electrostatic potential, so the relevant band gap energy is \( e\Delta \)). In this case the relevant band gap is that of STO if we assume no valence band offset for simplicity.

\(^9\) For the analogous discussion for a superlattice geometry see \([41]\).
carrier density. Essentially we are neglecting the kinetic energy of the electrons in the 2DEG by assuming the limit of a large density of states associated with the gas, and thus a negligible increase in the Fermi level with growing carrier density. This has been argued and used in the past in similar contexts (see e.g. [74, 75]), and is also confirmed by first-principles calculations that properly include the finite density of states at the interface (see e.g. [70]).

The consideration of a finite density of states, however, does affect the results by reducing the carrier density as a function of thickness. Such reduction is zero at \( d_c^{(p)} \) but grows with \( d \). This is an important point, considering (see section 3.3) that the carrier densities observed experimentally are substantially lower than expected. It is unlikely, however, that the aforementioned effect accounts for the whole of the discrepancy. In fact, at the LAO–STO interface, first-principles calculations which more properly include these effects produce a sheet carrier density comparable to that predicted by the initial model neglecting the kinetic energy of the 2DEG (see e.g. [70]).

The extension is simple. Equation (10) becomes

\[
\varepsilon_{\text{LAO}}^{(eq)} = \frac{\Delta + \sigma^{(eq)} / g}{d} = \frac{\sigma_c - \sigma^{(eq)}}{\epsilon},
\]

for a finite density of states \( g \) for the 2DEG (assuming parabolic 2D bands, the density of states is a constant, independent of energy). Solving for \( \sigma^{(eq)} \) gives

\[
\sigma^{(eq)} = \frac{\sigma_c d - \epsilon \Delta}{\epsilon / g + d}. \tag{13}
\]

The critical thickness remains the same, but the curve is depressed, tending to the same asymptotic value at large thickness \( \sigma^{(eq)} \rightarrow \sigma_c \), but the thickness for which \( \sigma^{(eq)} = \sigma_c / 2 \) is larger by \( \delta d = \epsilon / g \).

Reintroducing the value of \( \sigma^{(eq)} \) into equation (12), the thickness dependence of the field is obtained as

\[
\varepsilon_{\text{LAO}} = \frac{\sigma_c + g \Delta}{\epsilon + g d}, \tag{14}
\]

connecting with the model and analysis of [41]. It is no longer pinned to \( \Delta / d \).

We introduced \( g \) in equation (12) as the density of states of the 2DEG. This assumes that the holes at the surface are not mobile. If that is not the case, and we do have a 2DEG at the interface and a 2D hole gas at the surface, the model is still valid as is, taking \( g \) as the reduced density of states

\[
g = \frac{g_e g_h}{g_e + g_h} \tag{15}
\]

where \( g_e \) and \( g_h \) are the densities of states for electrons and holes, respectively. That is, \( g \) is the density of states corresponding to the reduced mass of electrons and holes.

Figure 4. Schematic band diagram of an interface between a polar thin film and a thick non-polar substrate, adapted from [71]. (a) The pristine system under the critical film thickness. (b) Electronic reconstruction. (c) The situation for the thick film limit (interface and surface well separated). (d) The introduction of a donor state at the surface and the subsequent electron transfer to the interface, reducing the electric field in the film. The dotted horizontal line represents the Fermi energy.

This process is identical to a Zener breakdown, although Zener tunnelling itself is not needed. By substituting the expression for the electric field in equation (7) and rearranging, it is simple to show [41, 73] that this transfer happens at a critical thickness for this pristine system

\[
d_c^{(p)} = \epsilon \Delta / \sigma_c, \tag{9}
\]

where \( \epsilon \) is the static dielectric constant of the film material (LAO). Beyond this thickness \( (d > d_c^{(p)}) \) the electron and hole carrier densities reach equilibrium, \( \sigma^{(eq)} \), pinning the potential drop to \( \Delta \), altering the electric field across LAO,

\[
\varepsilon_{\text{LAO}} = \Delta / d = (\sigma_c - \sigma^{(eq)}) / \epsilon. \tag{10}
\]

Rearranging this equation leads to the familiar expression for the equilibrium carrier density,

\[
\sigma^{(eq)} = \sigma_c - \frac{\epsilon \Delta}{d} = \sigma_c (1 - d_c^{(p)} / d). \tag{11}
\]

While the critical thickness is not affected, we must note here that the last two equations assume the strict pinning of the voltage drop to the value of the gap \( \Delta \) independently of the
3.2. Model reformulation

This model can be reformulated by quantifying the formation energy $\Omega_c$ of an area density $\sigma$ of electron–hole pairs as they are separated by the film’s electric field. The reason for this reformulation will become clear in section 4 when considering the additional effects of redox processes. We account for the competition between two terms; the energy cost associated with the formation of an electron–hole pair in the absence of the field, and the energy released from the charge compensation of the chemical charge by the transferred carriers (energy released by partly discharging a capacitor), as follows:

$$\Omega_c(\sigma) = \left( \sigma \Delta + \frac{1}{2} \frac{\sigma^2}{g} \right) - \frac{d}{2e} \left[ \sigma_{c}^2 - (\sigma_{c} - \sigma)^2 \right].$$  \hspace{1cm} (16)

The energy is taken with respect to $\sigma = 0$\(^{10}\).

The equilibrium carrier density $\sigma^{(eq)}$ is then obtained from minimizing $\Omega_c$ with respect to $\sigma$, with the condition that $\sigma \geq 0$,

$$\frac{\partial \Omega_c}{\partial \sigma} = \Delta + \frac{\sigma}{g} + \frac{d}{e} (\sigma - \sigma_{c}) = 0,$$  \hspace{1cm} (17)

which recovers equation (13), as long as $d$ is beyond a critical thickness (when $\sigma^{(eq)} = 0$) for this pristine system,

$$d_{c}^{(p)} = \frac{e \Delta}{\sigma_{c}}.$$  \hspace{1cm} (18)

recovering the usual expression (equation (9)).

For simplicity, in the remainder of this paper (unless stated otherwise) we will consider the limit of large $g$, thus also recovering the usual expression for $\sigma^{(eq)}$, (equation (11))

$$\sigma^{(eq)} = \sigma_{c} - \frac{\epsilon \Delta}{d}.$$  \hspace{1cm} (19)

As mentioned in the introduction, the concepts explained and the model here refer to the final equilibrium situation, not to the process of electrons tunnelling from the surface to the interface. Such a process can be quite involved since the situation evolves while the film is being grown, and it is clearly beyond the scope of this review. It is clear, however, that it would be naive visualizing the ideal process of having grown a film above the critical thickness and then having the electrons tunnel through at a later stage. It is quite healthy therefore to consider the point of view of Janotti et al\(^{[26]}\), which, coming from the field of semiconductor heterostructures, starts from the standpoint that an isolated neutral interface (or surface) of the kind that we have here has free carriers compensating for $\sigma_{c}$. Consider an $n$ interface and a $p$ surface coming together from an infinite distance (film thickness). The carriers in both (electrons at the interface, holes at the surface), starting from the perfectly screened situation, as in figure 4(c), will start to annihilate, equilibrating the Fermi level as in figure 4(b). Eventually, for enough proximity (the critical thickness), the annihilation is complete. The electric field across the film

\(^{10}\)A mean-field (Hartree) electron–electron term is not explicitly included because the mean-field electrostatic interaction is already included in the second term.

appears when the carriers start annihilating, increasing as the thickness is reduced until it saturates when all carriers have disappeared at (and below) the critical thickness. The two viewpoints are equivalent as regards the state that they describe for any thickness, although the rationales referring to the processes are opposite. The model and equations above are equally valid for these viewpoints.

Before we move on from this discussion, we note that the models formulated within this review consider a bare LAO thin film on an STO substrate under open-circuit electrical boundary conditions. The effect of a capping metal electrode on top of LAO has been examined from first principles under both open-circuit\([23, 76]\) and finite-$D$\([77]\) electrical boundary conditions, and experimentally under closed-circuit and finite-$V$ electrical boundary conditions\([78]\). We simply note here that clearly the picture is altered for these geometries and electrical conditions, with the top electrode providing an alternative source of electrons. Further discussion is beyond the scope of this review, and we refer the interested reader to\([23, 76–78]\).

3.3. Support for the electronic reconstruction idea and otherwise

With equation (9) one can therefore estimate the expected critical thickness for the electronic reconstruction in the LAO–STO system simply from the chemical interface charge of $0.5e/\Theta$, the dielectric constant of the LAO film and the relevant band gap of the system—the band gap of STO plus the interface valence band offset, and possible acceptor/donor states at the interface/surface. To convert this estimated critical thickness to units of LAO unit cells, we also need the out-of-plane LAO lattice parameter when strained to STO. Several values have been quoted in the literature for each of these properties, perhaps the most sensitive value being the LAO dielectric constant which has been suggested to be dependent on strain and electric field. If calculated $ab initio$, the dielectric constant is also dependent on the pseudopotential, basis set and functional. A recent review paper has summarized this sensitivity\([20]\), and the resulting variation in estimated critical thicknesses amounts to anything from 4 unit cells\([70, 72, 79]\) to 6 unit cells or even greater\([67, 80]\) being reported.

Perhaps the strongest support for the electronic reconstruction model is the fact that experimentally a critical thickness of 4 unit cells of LAO has been consistently reproduced from transport measurements with films grown by both pulsed laser deposition (PLD)\([7]\) and molecular beam epitaxy (MBE)\([35]\), once STO oxygen vacancies are removed by growth in sufficiently high oxygen partial pressure and annealing in oxygen\([6, 81]\). Consistent with these findings is the observation of a critical thickness of 4 unit cells of LAO on top of LAO has been examined from first principles\([82]\). A quite convincing recent study\([27]\) showed that, in agreement with the electronic reconstruction model, the critical thickness can be tuned by altering the chemical interface charge, $\sigma_{c}$, through an LAO solid solution with a non-polar cubic oxide (STO$_{1−\chi}$LAO$_{\chi}$). For such a system the formal charges argument gives $\sigma_{c} = (x/2) e/\Theta$. 

and from equation (9) the electronic reconstruction model would predict a $1/x$ dependence on the critical thickness $d_c^{(p)}$ of the diluted film, which is indeed what was obtained in the experiment [27]$^{11}$.

This argument would seem to contradict the discussion above that led to the $\sigma_c = (1/2) e/\Theta$, result as being protected by symmetry. Assuming random alloying of both the A (Sr and La) and the B cations (Ti and Al) through the film, in an average (virtual crystal approximation) view, the system would remain centrosymmetric and thus with $P = P_0/2$. The paradox is however resolved by noting that within the average-alloy scenario contemplated, the relevant quantum of charge renormalizes to $x e$, and $\sigma_c$ is still $P_0/2 = (x/2) e/\Theta$, consistent with both the symmetry arguments above and the experiment interpretation of [27].

Additional experimental support for the electronic reconstruction model includes the indirect observation of a built-in electric field in LAO before the critical thickness through the electrostriction effect [70], and structural [68] and tunnelling measurements [78, 83]. However it should be noted that with spectroscopic measurements, researchers have not been able to detect a sizably varying field with thickness [32, 35, 36].

Finally further strong support for the electronic reconstruction model is provided by the fact that several other polar interfaces are also conducting, such as interfaces between SrTiO$_3$ and LaTiO$_3$ [84], LaGaO$_3$ [85], LaVO$_3$ [86], KTaO$_3$ [87], KNbO$_3$ [87], NaNbO$_3$ [87] and GdTiO$_3$ [88].

Not all experimental observations are in direct agreement with the electronic reconstruction model, however. Measured electron carrier densities are usually found to be approximately one order of magnitude lower than that predicted from equation (11) using $\sigma_c = (1/2) e/\Theta$ [7]. The discovery of a sizable density of (trapped) Ti 3d-like states below the critical thickness (at just 2 unit cells of LAO) [28–30, 32] with core-level spectroscopic measurements suggests that the breakdown may occur almost immediately. Additionally no surface hole carriers have been measured from transport [7] except when an STO capping layer is added [89, 90], and no surface hole states have been seen near the Fermi level [33, 34]. These last findings, however, can be explained if the surface states are immobile [80, 91]. Probably more significant, however, is the apparent disappearance of conductivity at any LAO thickness for samples grown at very high oxygen partial pressures [37, 38]. This is a puzzling result. We can only speculate that this might be connected with the dramatically reduced carrier densities [92–96] (and degree of hydroxylation [97]) observed at certain La:Al off-stoichiometries, which will be further discussed in section 5. Other possible mechanisms proposed on the basis of interface effects will be considered in that section too.

In the next section we explore the possibility of an alternative source of carriers, namely the creation of surface redox defects, which can alleviate the polar catastrophe and explain some of the issues mentioned within this section with the purely electronic reconstruction mechanism. We model the creation of these processes and discuss them as the possible origin of the LAO–STO 2DEG in the light of available experimental data.

4. Redox screening

4.1. The general model

We now introduce the possibility of there being surface redox reactions that leave immobile surface defects with a net charge, and free carriers that can move to the interface (see figure 4(c)). A prototypical example [8, 71, 98–100] would be that of oxygen vacancies at the surface, whereby

$$\text{O}^{2-} \rightarrow \frac{1}{2}\text{O}_2 + V_O + 2e^-.$$  \hspace{1cm} (20)

Since the system with the original $\text{O}^{2-}$ was neutral and the two electron carriers are free to go, the vacancy, $V_O$, is then left with an effective charge $Q = +e$. If the surface is in contact with the atmosphere and water is present, one can also think of the redox hydroxylation reaction

$$2\text{O}^{2-} + \text{H}_2\text{O} \rightarrow \frac{1}{2}\text{O}_2 + 2\text{OH}^- + 2e^-,$$  \hspace{1cm} (21)

in which two defects are formed (the two OH$^-$ groups) and two carriers are generated (note that the non-redox hydroxylation process, $\text{H}_2\text{O} \rightarrow \text{OH}^- + \text{H}^+$, does not a priori affect the electrostatics across the film, except through small alterations in the surface valence band maximum and subsequent closing of the effective gap, $\Delta$ [101]). Each centre now has $Q = +e$. The mechanism then generalizes to the formation of redox defects with $Q = Ze$ at the surface and a number $|Z|$ of free carriers at the interface. If $Z > 0$, the interface carriers will be electrons and the surface defects donors. $Z = 1$ represents single donors (e.g. the hydroxyl groups above), $Z = 2$, double donors (the O vacancies). This corresponds to an oxidation of the surface. The opposite is also possible: a surface reduction, with $Z < 0$, in which case the surface redox defects will be acceptors, and the carriers at the interface, holes, which would be the stable (favourable) situation in the case of a p interface. These redox processes are not dissimilar to the pure electronic reconstruction that we had in section 3 (electron–hole pair formation) except that now the surface carriers are immobile (surface defect) and have a charge $Q = Ze$, while the interface carriers are mobile electrons or holes. We consider $Z > 0$ in the figures and examples in what follows.

Let us extend the initial model (equation (16)), allowing for the formation energy of an area density $n$ of such redox defects [71]. These defects compete with electron–hole pairs, with density $\sigma$, from the electron-transfer mechanism:

$$\Omega(n, \sigma) = Cn + \frac{1}{2}an^2 + \sigma \Delta$$

$$+ \frac{d}{2\epsilon} \left[ (\sigma_c - nQ - \sigma)^2 - \sigma_c^2 \right].$$  \hspace{1cm} (22)

This now incorporates the formation energy of an isolated redox defect in the absence of a field, $C$, and a defect–defect interaction term, accounted for in a mean field as $\frac{1}{2}an^2$. As in the previous section, the mean-field electrostatics is
captured by the capacitor discharging term (the last term), and therefore the \( \frac{1}{2} \alpha n^2 \) term describes the mean-field defect–defect interaction beyond electrostatics (e.g. strain, chemical). The capacitor discharging term now incorporates the compensation due to both charges: the electron–hole pairs (\( \alpha \)) and the ones coming from the redox processes (\( nQ \)). The free charge density at the interface is now the sum of the two, \( \sigma_{2DEG} = \alpha + nQ \).

The condition \( \sigma \geq 0 \) implies

\[
\sigma^{(eq)} = \sigma_c - nQ - \frac{\epsilon \Delta}{d}.
\]

The condition \( \sigma \geq 0 \) implies

\[
nQ \leq \sigma_c - \frac{\epsilon \Delta}{d},
\]

meaning that in the plot of \( nQ \) versus \( d \) (see figure 5) there is a separation of two regions, one with conventional electron–hole formation, the other without it. As expected, the boundary line hits the \( n = 0 \) line at \( d = d_c^{(p)} \), the critical thickness for the appearance of electron–hole pairs in the absence of redox defects. It is important to note here that the energy \( \Omega \) will be continuous across the boundary between regions, although its derivative will not.

Let us now find the equilibrium concentration of surface redox defects.

4.1.1. Region I: \( \sigma^{(eq)} = 0 \). Consider first the region in which there is no electronic reconstruction, \( \sigma^{(eq)} = 0 \). The energy becomes

\[
\Omega(n) = Cn + \frac{1}{2} \alpha n^2 + \frac{d}{2 \epsilon} \left[ (\sigma_c - nQ)^2 - \sigma_c^2 \right],
\]

and \( n^{(eq)} \) is found as ever from

\[
\frac{\partial \Omega}{\partial n} = 0 \Rightarrow n^{(eq)} = \frac{Qd \sigma_c - \epsilon C}{Q^2 d + \epsilon \alpha},
\]

as already presented in [71]. In this case, there is a critical thickness for the onset of redox screening, i.e., when \( n^{(eq)} \geq 0 \), which happens at the critical thickness

\[
d_c = \frac{\epsilon C}{Q \sigma_c}
\]

which is relevant only while in the \( \sigma = 0 \) regime. That means that \( d_c \) is meaningful if it is smaller than \( d_c^{(p)} \). Using the definitions of [71], this happens when \( C \leq Q \Delta \). In practice, this corresponds to cases where the redox process is energetically more favoured than the creation of electron–hole pairs in zero field.

The \( \alpha \) parameter is expected to be small [71]; therefore let us see what happens for \( \alpha = 0 \). Then

\[
n^{(eq)} = \sigma_c - \frac{\epsilon C}{Qd},
\]

which, if compared with equation (25), shows that the equilibrium redox defect density runs parallel to the boundary between the two regions (\( \sigma^{(eq)} = 0 \) and \( \sigma^{(eq)} > 0 \), asymptotically as \( 1/d \), towards \( \sigma_c \).

For \( \alpha > 0 \), the \( n^{(eq)} \) onset is still the same, at \( d_c \), and the curve still tends asymptotically to \( \sigma_c \) as \( 1/d \). However, the behaviour in between is not as \( 1/d \), and the value of \( \alpha \) pushes the curve down and makes it cross the boundary at some thickness \( d_x \). Such crossing can be obtained by finding the thickness for which \( n^{(eq)} \) equals the boundary density, equation (25),

\[
n^{(eq)} = \sigma_c - \frac{\epsilon \Delta}{Q d_x},
\]

which, using equation (27), gives

\[
d_x = \frac{\epsilon \Delta \alpha}{QC - Q^2 \Delta + \alpha \sigma_c}.
\]

This last equation becomes clearer when written as follows:

\[
d_x = d_c^{(p)} \frac{d_a}{d_a - \delta d},
\]

where

\[
d_a \equiv \frac{\epsilon}{Q^2} \alpha \quad \text{and} \quad \delta d \equiv d_c^{(p)} - d_c.
\]
For the crossing we seek, \( d_e > d_c^{(p)} \), we then need that \( d_a > d_a - 3 \delta d \) (which happens whenever \( d_e < d_c^{(p)} \)), and we need the denominator to be positive, i.e.,

\[
d_a > 3 \delta d.
\]

This means that if the onset for redox defect stabilization happens before the onset for electron–hole pair formation, the equilibrium line will remain in region I (no additional electron–hole formation, \( \sigma = 0 \)) for all thicknesses, unless the surface defect mutual repulsion is strong enough. In this case of \( \sigma = 0 \) the free charge at the interface is simply \( n^{(eq)} Q \) and can reach \( \sigma_c \) rather quickly. It is interesting to briefly note here a comparison with the experimental results in [70], where complete screening of the electric field was inferred through the electrostriction effect for films as thin as six LAO perovskite layers.

It remains to be seen what happens if the equilibrium line crosses to region II, or if the onset had already occurred in region II, or, even if there is an equilibrium line always in region I, whether there may be another in region II. Before we get there, let us find the value of \( n^{(eq)} \) at the crossing point. For that we introduce the value of \( d_e \) of equation (31) into \( n^{(eq)} \) as in equation (30),

\[
n^{(eq)}(d_e) = \frac{c_e}{Q} \left[ \frac{\Delta}{Q} - \frac{\epsilon \Delta q C - Q^2 \Delta + \alpha c_e}{\epsilon \alpha \Delta} \right],
\]

which yields

\[
n^{(eq)}(d_e) = \frac{Q \Delta - C}{\alpha}, \tag{36}
\]

which is below the asymptotic limit of \( c_e / Q \) if \( \delta d < d_a \), and hence consistent with the previous discussion leading to equation (34).

### 4.1.2 Region II: \( \sigma^{(eq)} > 0 \)

In this region the equilibrium electron–hole planar density \( \sigma^{(eq)} \) is the one given by equation (24) in the previous section. We now find the equilibrium density of redox defects by minimizing \( \Omega \) (in equation (22)) with respect to \( n \):

\[
\frac{\partial \Omega}{\partial n} = C + \alpha n - \frac{d}{\epsilon} (\sigma_c - n Q - \sigma) Q = 0. \tag{37}
\]

Introducing \( \sigma^{(eq)} \) and solving, we obtain

\[
n^{(eq)} = \frac{Q \Delta - C}{\alpha}, \tag{38}
\]

which is thickness independent and coincides with the boundary value of \( n^{(eq)}(d_e) \) coming from region I (see equation (36)). This means that if \( n^{(eq)} \) of region I crosses the boundary, it continues horizontally into region II. If, on the other hand it does not cross, then there is no minimum for \( n \) in region II (if \( d_a < \delta d \), then \( (Q \delta - C) / \alpha \) gives a value beyond \( \sigma_c / Q \)), and the complete solution found in region I remains unique.

Finally, if the onset of defect formation happens for thicker samples than the onset for electron transfer, i.e. \( d_e > d_c^{(p)} \), then the equilibrium \( n \) in region II becomes negative for all thicknesses, meaning that \( n^{(eq)} = 0 \) always, i.e. no redox defects appear. In this case we recover the solution of the previous section, \( \sigma_{2D} = \sigma \).

### 4.2. Comparison with first-principles calculations

#### 4.2.1. The formation energy of redox defects

The energy \( \Omega \) within sections 3 and 4 accounts for the formation energy of both redox defects and electron–hole pairs. Sometimes there is interest in the energies themselves (for instance, if comparing with electronic structure calculations). Depending on what is being compared, the relevant energy may be the energy of the system with redox defects referred to the energy that the system would have if only allowing for electronic transfer. In that case the relevant energy is

\[
\tilde{\Omega}(n) = \Omega(n, \sigma^{(eq)}) - \Omega_c(\sigma^{(eq)}). \tag{39}
\]

It does not affect the equilibrium concentrations for \( n \), since we are just adding an \( n \)-independent energy to what we had. The only difference in the analysis is that there would be a discontinuity in the derivative of the energy \( \tilde{\Omega} \) at \( d_c^{(p)} \) because of there being a kink in \( \Omega_c \). This is indicated in figure 5 by a vertical dotted line, effectively defining a third region. The resulting energy in the different regions is

\[
\tilde{\Omega}(n) = \begin{cases}
C n + \frac{1}{2} \epsilon \alpha n^2 + \frac{d}{\epsilon} \left[ (\sigma_c - n Q)^2 - \sigma_c^2 \right] & \text{for } d < d_c^{(p)} \\
C n + \frac{1}{2} \epsilon \alpha n^2 + \frac{d}{\epsilon} \left[ (\sigma_c - n Q)^2 - \sigma_c^2 \right] & \text{for } n Q \geq \sigma_c
\end{cases} \tag{40}
\]

#### 4.2.2. The formation energy of one redox defect

Given the fact that \( \Omega(n) \) is the energy for a given density of defects \( n \), the formation energy of one redox defect in an infinite system with a defect density \( \rho \) would be

\[
E^D(n) = \frac{\partial \tilde{\Omega}(n)}{\partial n}. \tag{41}
\]

However, to compute the formation energy of one redox defect from first principles one would calculate the difference in energy between a thin film with one redox defect in a given supercell, which defines the concentration \( n \), and the energy of the same supercell without the defect. That quantity we will call \( E^D \), and it relates to the previous magnitudes in the following way:

\[
E^D(n) = \frac{\tilde{\Omega}(n)}{n} = \frac{1}{n} \int_0^n E^D(n') \, dn', \tag{42}
\]

which results in

\[
E^D(n) = \begin{cases}
C n + \frac{1}{2} \epsilon \alpha n^2 + \frac{d}{\epsilon} \left[ (\sigma_c - n Q)^2 - \sigma_c^2 \right] & \text{for } d < d_c^{(p)} \\
C n + \frac{1}{2} \epsilon \alpha n^2 + \frac{d}{\epsilon} \left[ (\sigma_c - n Q)^2 - \sigma_c^2 \right] & \text{for } n Q \geq \sigma_c
\end{cases} \tag{43}
\]
Figure 6. Redox defect formation energy (per defect) \( E_f \) as a function of the thickness \( d \), where it is compared with several surface vacancy formation energies obtained from first principles [98, 99]. Figure 6 shows the \( E_f \) for various concentrations, as a function of the thickness \( d \), where it is compared with several surface vacancy formation energies obtained from first principles [98, 99]. The agreement is remarkable, taking into account that the parameters could be independently computed (C is hard to pin down in comparisons with experiments since it contains the chemical potential, which is hard to establish; but this is easier in comparison with calculations where the energy of the reference structures is well defined; see the appendix in [71]). The formation of surface oxygen vacancies in LAO-STO has been further ratified within the careful calculations of [104] only for the ideal concentration of oxygen vacancies, without considering oxygen vacancy concentration dependence.

4.3. Generalization

The situation described in this section can be seen as that of competing screening mechanisms driven by the depolarizing field, both \( n \) and \( \sigma \) being positive definite, and both with their respective energy thresholds. One can think of the general screening of a thin film by different such mechanisms by rewriting equation (22) as

\[
\Omega((\sigma_i)) = \sum_i (\sigma_i \Delta_i + \frac{1}{2} \sigma_i \sigma_i^2) + \frac{d}{2\epsilon} \left[ (\sigma_c - \sum_i \sigma_i)^2 - \sigma_c^2 \right].
\] (44)

In the case of electronic reconstruction the quadratic term can also be included, as \( \omega_i = 1/g_i \), the inverse density of states, accounting for the kinetic energy of the carriers, as discussed in sections 3.1 and 3.2.

Similarly, this generalized energy could include bilinear couplings in the different variables, of the form \( \sum_{ij} \beta_{ij} \sigma_i \sigma_j \). In the case of \( \sigma \) and \( nQ \), if the energy in equation (22) had an added term \( \beta nQ \sigma \), the equilibrium electron density (equations (24)) would read

\[
\sigma^{(eq)} = \sigma_c - nQ \left( 1 + \frac{\beta \epsilon}{d} \right) - \frac{\epsilon \Delta}{d}
\] (45)

and the phase boundary (in equation (25)) given by the condition \( \sigma > 0 \) would transform into

\[
nQ < \frac{\sigma_c - \frac{\epsilon \Delta}{d}}{1 + \frac{\beta \epsilon}{d}}.
\] (46)

Such a term would describe the change in effective gap when in the presence of surface carriers due to the electronic reconstruction. It is important to state now, however, that under such interactions it is possible that the relevant defects would cluster, an effect beyond the mean-field description implied in the equations. This discussion will not be pursued any further here.

4.4. Experimental signatures and discussion

We conclude this section with a summary of experimental evidence for and against the redox screening mechanism. We concentrate on the LAO–STO system for which there are numerous studies for comparison.

The first expectation from redox screening is obviously that the surface remains insulating (except for ionic diffusion) since it is charge compensated by bound charge centres (e.g. oxygen vacancies), whilst the interface is screened by electrons (which may or may not be mobile). As already mentioned, hole conduction at the LAO surface has not been found for any thickness of LAO, whilst electronic conduction appears at the LAO–STO interface at 4 unit cells of LAO or more [7]. Ti 3d-like states (Ti 3+) have been observed from spectroscopic measurements even below the critical thickness of 4 unit cells, and for sizes as thin as just 1 unit cell [28–30, 32]. This suggests the possibility of the redox screening mechanism appearing almost immediately with LAO growth, since with the purely electronic reconstruction one would expect surface hole conduction and only an electronic transfer to the interface once LAO is approximately 4–6 unit cells thick (see section 3). It is indeed shown in [71] following the same recipe as in section 4 that the formation of surface oxygen vacancies in LAO–STO is likely to occur for LAO films as thin as 1–3 unit cells thick.

The rapidly increasing interface electron density with film thickness predicted within the redox model (for a certain range of \( \sigma \)) is consistent with the experimental results in [70], where nearly complete screening of the electric field was inferred through the electrostriction effect for films as thin as six LAO perovskite layers. This is too thin for being explained by the electronic reconstruction model alone.

If redox formation is indeed responsible for the observed Ti 3d states, the next question is why for films thinner than 4 unit cells these electrons are not available for conduction. We briefly mention three possibilities that come to mind. Firstly it has been suggested that the 2DEG lies in several Ti 3d sub-bands, some of which are not mobile due to Anderson localization [105]. The other two mechanisms are related to the experimental observation that the immobile Ti 3d levels are ‘in-gap’ states [33, 106–108] at higher binding energies. The second possibility is that acceptor defect levels are formed during the LAO growth process, such as...
from cation intermixing (see section 5). Cation intermixing has been readily observed in most LAO–STO samples [25, 68, 109–115].

The third possibility is that such states could be formed from the point charge of the surface redox defects which generate trapping potentials, as discussed in [71]. As the LAO film thickness increases the traps change from deep and few to shallow and overlapping. Hence a transition from insulating to conducting occurs at larger LAO film thickness than the redox formation itself. The character of the transition would depend on the charge of the redox defects $Z$. For even $Z$ (as in the double-donor case of the oxygen vacancies), the transition would be that of band overlap, whenever the band of filled states would touch the bottom of the conduction band. For odd $Z$ (as in the case of OH$^-$ defects) the partly filled gap states would undergo an insulator to metal transition of the Mott–Anderson type [71].

Several other recent experimental findings, some of which are incompatible with the electronic reconstruction, find natural explanation within the redox screening model. XPS measurements have revealed very little, if any, evidence for core-level broadening or shift with increasing film thickness [32, 35, 36], at odds with what one might expect from a constant electric field within LAO before the electronic reconstruction critical thickness. Within the redox screening model, it can be trivially shown that the potential drop across the LAO film is essentially independent of the thickness (if $\alpha$ is small) [71]. Using the parameters for LAO–STO the potential drop for each LAO layer added was predicted to be between 0.0 and 0.2 eV [71], much smaller than within the electronic reconstruction model (approximately 0.7–0.9 eV). The predicted pinning of the potential drop (and hence reduction in electric field with thickness) is also consistent with the reduced cation–anion buckling with increasing LAO thickness as observed using SXRD [68].

Finally we note the observation that the field effect switching of the LAO–STO conductivity [7, 8] is intimately linked with surface charge writing [116, 117] and that the process is dependent on the existence of water in the surface environment [118]. These experiments are probably linked with electrochemical processes [119, 120], and are all compatible with the redox screening model. Applying a biased tip to the surface alters the field across the LAO film which either increases or decreases the stability of surface redox defects (and hence interface charge carriers and surface charge) depending on the sign of the bias. Such electrochemical processes have also been predicted during the switching of ferroelectric thin films [121–125].

5. Interface effects

So far we have assumed perfect interfaces and thin two-dimensional charge distributions. In this section we explore the effect on the phenomenology of deviations from such assumptions.

5.1. The interface dipole due to the electrons

The free carriers at the interface do screen (some of) the chemical charge at the interface, but their quantum spread in $z$ generates an interface dipole that has been proposed to alter the relative alignments between LAO surface and STO’s bottom of the conduction band. This effect is well described in the literature (see e.g. [126] and references therein). The relevant alignment [126] is the one given by the filling of the carrier bands in the 2DEG, as illustrated in figure 7. It depends on the zero-point energy of the band, and therefore the perpendicular effective mass of the carriers, and on the 2DEG band filling, which depends on the carrier density and the parallel effective mass of the carriers. Discussions on the character of the 2DEG bands and corresponding effective masses are beyond the scope of this review.

5.2. Interface cation intermixing

Both during and after sample growth there is the possibility of inter-diffusion of atoms across the interface. Such processes have attracted attention lately [25, 68, 109–114] as one of the possible explanations for the origin of 2D conduction in LAO/STO interfaces [21]. The argument used is simply that, since La is a known dopant (donor) in STO, La cations diffusing from LAO into STO could dope the interface and thus make it conducting.

Let us consider here what to expect from ions exchanging across the interface\textsuperscript{12}. This discussion will focus on the effects of having a given inter-diffusion profile in a sample, rather than the energetics and kinetics that might give rise to it, since equilibrium is not expected to be reached in this aspect. We shall concentrate here on inter-diffusion situations starting from the pristine stoichiometric interface, and all the exchanges happening within a region around the interface whose width is smaller than other length scales in the problem. The effects on deviations from stoichiometry, and of diffusion processes from/to reservoirs extrinsic to the interface, are qualitatively different and will be treated briefly in section 5.3.

\textsuperscript{12} In what follows we will make two approximations to simplify discussions, both of which can be trivially added to the models. Firstly we take the pristine interface dipole to be zero, and secondly we make the approximation that the dielectric constant is unaltered by inter-diffusion across the interface.
5.2.1. Interface doping generated by inter-diffusion. On the basis of chemical (steric) arguments and on quite a wealth of accumulated knowledge on perovskites, the most likely inter-diffusion processes by far are swaps of like atoms, namely, B cations (Al$^{3+}$ cations in LAO substituted by Ti$^{4+}$, and Ti$^{4+}$ in STO substituted by Al$^{3+}$) and A cations (La$^{3+}$ to Sr$^{2+}$ on one side, Sr$^{2+}$ to La$^{3+}$ on the other). The swapping of O$^{2-}$ anions across the interface is also likely, but of no consequence.

Each cation substitution gives rise to a net bound charge plus a compensating carrier. La$^{3+}$ replacing Sr$^{2+}$ gives a $+e$ centre with an associated electron, and thus the donor mentioned in the previous subsection, and so do the other substitutions with obvious signs for their charges.

The first important point to make here is that such inter-diffusion should not be expected to dope the interface. Be it A or B cations swapping across the interface, there will be as many ‘wrong’ cations on one side as on the other, and thus as many donors on one side as acceptors on the other. The electrons of the donors annihilate with the holes of acceptors, as long as both dopants are located in the proximity of the interface, the situation considered here. Similarly, there will be as many Ti$^{4+}$ donors in LAO as Al$^{3+}$ acceptors in STO. The inter-diffusion will then give rise to a distribution of fairly immobile charges on both sides of the interface, with zero total charge, since there are as many charges $+e$ as $-e$. Indeed it is well known that the LAO–STO solid solution is insulating.

While intermixing is not expected to dope the interface, it can have implications for the transport properties and critical thickness of an interface 2DEG appearing by means of another mechanism. This is likely since the intermixing can produce the aforementioned trapping acceptor states lying below the conduction band of the pristine interface.

5.2.2. The electric field, interface dipole and interface shift. The second important point to make here is that such inter-diffusion does affect the electrostatics, but only in the interface region. It does not affect the electric field emanating from it into the film. This is a direct consequence of Gauss’s law: the net flux of the electric displacement field across surfaces parallel to the interface equals the net charge contained in the region within them. In the absence of charges diffusing in from other regions, the net charge per unit area is still $\sigma_{c}$, and the field emanating from the interface into the film is again $\mathbf{E} = \sigma_{c}/\epsilon_{\text{LAO}}$ (the field in the substrate being zero).

The main effect of cation interdiffusion across the interface is illustrated in figure 8: the electrostatic potential is shifted with respect to the one arising from a pristine interface. It is essentially the effect of the interface dipole generated by the swap of charges with its characteristic potential drop $\Delta V$ (or rise). Alternatively, this effect can be interpreted as an effective shift of the interface position, $\Delta z$ in the figure. This can be a productive way of seeing it, renormalizing the film thickness in the analysis in sections 1–4.

![Figure 8. Electrostatics near an interface with $\sigma_{c}$ net charge per unit area, around which there is a region with cation intermixing (dark grey band). Electrostatic potential energy $V$ versus distance perpendicular to the interface $z$. The boundary condition is taken as a constant potential on the side of the substrate. The slope of the outgoing potential to the right is $\sigma_{c}/\epsilon$ regardless of the inter-diffusion, the latter producing a potential shift $\Delta V$, or an apparent shift in interface position, $\Delta z$.](image)

5.2.3. The sign of the dipole and shift. The third point in this section is on the sign of such an effect. Figure 9 shows quantitative examples for the behaviour of the electrostatic potential energy for electrons in two cases of inter-diffusion, one for B cation exchange the other for A. The former, having Al and Ti ions exchanged, produces a voltage drop, or a right shift of the effective interface position, and thus renormalize the film thickness to a smaller value. The latter produces the opposite. From a thermodynamic point of view, B cation inter-diffusion will be favoured by the electrostatics, whilst A cation inter-diffusion will tend to be suppressed$^{13}$.

5.2.4. The scale of the interface shift. The fourth point is on the scale of interface shifts. It is obvious that they will only be noticeable if on the scale of $\sigma_{c}$, i.e. if inter-diffusion is sizable compared with 50% per cation layer (inter-diffusion on the ppm scale can be happily ignored). Let us obtain a quantitative estimate for $\Delta z$.

Let us consider the $(x, y)$-averaged charge density associated with inter-diffusion as $\delta_{L}(z)$ and $\delta_{R}(z)$ for the charge at the left and the right sides of the interface, respectively—the former with support for $z < 0$ only, the latter, for $z > 0$ (we consider the pristine interface at $z = 0$). The fact that there are no external charges contributing to these densities is reflected in

$$\sigma_{d} = -\int_{-\infty}^{0}\delta_{L}(z)\,dz = \int_{0}^{\infty}\delta_{R}(z)\,dz.$$  \hspace{1cm} (47)

This description allows for A or B cation intermixing, and for both. $\sigma_{d} > 0$ implies more B cation exchanges, as in the

$^{13}$ Remember that these arguments are based on the original assumption of zero pristine interface dipole. A large non-zero pristine dipole of appropriate sign can reverse this trend.
upper panel of figure 9, while A cation exchanges dominate for negative $\sigma_d$.

The electric displacement at any $z$ position can be written as

$$D(z) = D(-\infty) + \int_{-\infty}^{z} \delta(z') \, dz'.$$  \hfill (48)

The zero-field substrate boundary condition gives $D(-\infty) = 0$. Assuming linear dielectric response, the electric field anywhere is given by $D(z) = \epsilon_0 \varepsilon(z) + P(z) = \epsilon(z) \, \mathcal{E}(z)$, and thus, the field for any negative $z$ is given by

$$\mathcal{E}(z) = \frac{1}{\epsilon_{\text{STO}}} \int_{-\infty}^{z} \delta_L(z') \, dz'.$$  \hfill (49)

The electrostatic potential for anywhere in the substrate ($z < 0$) is

$$\psi(z) = \psi(-\infty) - \int_{-\infty}^{z} \mathcal{E}(z') \, dz'.$$  \hfill (50)

Defining the reference potential as $\psi(-\infty) = 0$, and using the previous expression for the field, we get (still for $z < 0$)

$$\psi(z) = -\frac{1}{\epsilon_{\text{STO}}} \int_{-\infty}^{z} dz' \int_{-\infty}^{z'} \delta(z'') \, dz''.$$  \hfill (51)

which, reordering the double integral, gives

$$\psi(z) = -\frac{1}{\epsilon_{\text{STO}}} \int_{-\infty}^{z} \delta(z'') \, dz'' \int_{-\infty}^{z} dz'.$$  \hfill (52)

$$= -\frac{1}{\epsilon_{\text{STO}}} \int_{-\infty}^{z} \mathcal{E}(z) \, dz.$$  \hfill (53)

$$= -\frac{1}{\epsilon_{\text{STO}}} \sigma_z \, (z - \langle z \rangle_z),$$  \hfill (54)

where $\sigma_z$ is the total charge integrated up to $z$, and $\langle z \rangle_z$ is the average charge position for the charge distribution up to $z$. The total potential change within the substrate up to the interface is thus

$$\psi(0^-) = -\frac{\langle z \rangle_{\text{STO}}}{\epsilon_{\text{STO}}},$$  \hfill (55)

$\langle z \rangle_{\text{STO}}$ indicating the average position of the diffusion charge in STO, and $\sigma_d$ as defined in equation (47) (taking care of the sign).

Within the LAO side the potential has two additive components, namely, as originated by $\sigma_c$ and $\sigma_d$. The first one gives the straight line starting at $z = 0$ in figure 9, the second one saturates to a constant once $\delta(z)$ dies out (the sum of the two giving the potential curves in the figure). The shift of potential due to $\sigma_d$ in LAO is found analogously to that for STO, giving the total potential saturating at

$$\psi(\infty) = \sigma_d \left( \frac{\langle z \rangle_{\text{STO}}}{\epsilon_{\text{STO}}} - \frac{\langle z \rangle_{\text{LAO}}}{\epsilon_{\text{LAO}}} \right).$$  \hfill (56)

The change in electrostatic potential energy for the electrons, as presented in figure 9, $V = -e \psi$, is thus

$$\Delta V = \sigma_d e \left( \frac{\langle z \rangle_{\text{LAO}}}{\epsilon_{\text{LAO}}} - \frac{\langle z \rangle_{\text{STO}}}{\epsilon_{\text{STO}}} \right).$$  \hfill (57)

This expression reflects the dipole originated by the exchange, allowing a very intuitive interpretation of such effects. Following the definition of the sign of $\sigma_d$, a cation exchange gives positive $\Delta V$, which is negative for B cation diffusion.

As illustrated in figure 8, this effect can be described as a shift in the apparent position of the interface, $\Delta z$. Given the slope of the right-outgoing potential energy, $e \sigma_c/e_{\text{LAO}}$, this gives

$$\Delta z = \frac{\sigma_d}{\sigma_c} \left( \frac{\langle z \rangle_{\text{LAO}}}{\epsilon_{\text{LAO}}} - \frac{\langle z \rangle_{\text{STO}}}{\epsilon_{\text{STO}}} \right).$$  \hfill (58)
The thickness dependence of film properties, including film critical thickness, should then refer to the effective interface position, resulting in a net shift in the thickness axis of figures like figure 5.

One consequence of equation (58) is that in a system like LAO/STO, with $\epsilon_{\text{STO}} \gg \epsilon_{\text{LAO}}$, the width of the cation distribution in STO hardly affects the final shift\(^{14}\),

$$\Delta z \approx \frac{\sigma_d}{\sigma_c} \langle z \rangle_{\text{LAO}}. \tag{59}$$

Considering now an LAO film of thickness $d$, this expression implies that the field across the film will be noticeably screened by cation inter-diffusion if (i) $\sigma_d \sim \sigma_c = +0.5e/\Theta$, i.e. a very substantial amount of Ti\(^{4+}\) is going into LAO, and (ii) $\Delta z \sim d$, i.e. Ti cations penetrate the whole film. Indeed, perfect screening would only be possible if half a monolayer of Ti cations all traversed the whole film to sit on the surface, while their Al counterparts would go to the STO side staying close to the interface. Such cases were confirmed to screen the field within the calculations of [104, 109, 110] and were obviously found to be quite stable and even exothermic. Whether these swaps across the whole film actually happen depends on the kinetics of the process.

5.3. Interfacial stoichiometry deviations

If in the growth process there is a cationic imbalance giving rise to a net deviation of the compositional charge, the effect is a trivial renormalization of $\sigma_c$. Unlike for the previous cases, where there were charge redistributions around the interface, with no change in the net charge, this section considers effects that alter the total interfacial charge.

Recently several studies have discussed the role of La:Al off-stoichiometries, which can be modulated by the laser parameters of the pulsed laser deposition, on the carrier density [92–96]. It was shown that when the La/Al ratio differs from unity by just a few per cent, the carrier density can drop by two orders of magnitude. It is clear that cation vacancies will affect $\sigma_c$, induce acceptor states and hence reduce the interface carrier density. A concentration of just 2% A site vacancies across the whole of a LAO film of 10 unit cells would correspond to a charge close to half a quantum, 0.5$e$/\Theta. Further discussion is beyond the scope of this paper and we refer the reader to more detailed studies [127, 128] of various kinds of non-stoichiometric LAO–STO systems.

These studies show that several kinds of defect structures are thermodynamically stable, for example the combined La vacancy at the LaO interface and O vacancy at the AlO$_2$ surface, which screen the internal electric field and in some cases remove the interface carriers. In fact, combined defects at surfaces and interfaces of this last kind represent a charge screening mechanism similar to the ones modelled within this review: instead of the scenario of surface O vacancies and interface free electron carriers, there are surface O vacancies

\(^{14}\) Remember that this prediction is under the original assumption that $\epsilon$ is unaltered by intermixing either side of the film, which is probably inaccurate [27]. Deviations from such assumptions can be trivially added to the model.

and interface cation vacancies. The equilibrium occurrence of such combined defects can be treated with the same model (the $C$ parameter would now include the combined formation energy of the two defects), but the key difference is that this mechanism does not produce carriers for the 2DEG.

It is likely that in several systems there also exists a total deficit of oxygen, via the creation of oxygen vacancies in the substrate during the growth process or by other means, even if they are not stabilized or favoured by the electrostatics. Here each vacancy, which is probably created near the substrate surface, creates two electrons which are loosely bound near the vacancy. These processes have been suggested to play a role in LAO–STO when grown under insufficient oxygen partial pressure [6]. They may also be created at the bare STO surface when subject to ultraviolet light irradiation [129] or vacuum cleaved [130].

The unexpected conductivity observed at the non-polar interface between an amorphous LAO film and a crystalline STO substrate [54, 131–134] has recently also been explained as the creation of oxygen vacancies in the STO substrate during the growth process [132, 134]. If the system is subjected to post-growth annealing in oxygen, which presumably removes these vacancies, the conductivity is no longer observed [134]. These processes are also possibly happening in other hetereopitaxial non-polar oxide systems [135].

Interestingly, alterations in the net charge can also arise if there are half-unit-cell steps at the interface, with both n and p interface terraces. In such a case there is a continuous variation of $\sigma_c$, from $-0.5e/\Theta$ to $+0.5e/\Theta$, depending on the ratio of interface area covered by each kind, including a total removal of the polarization discontinuity for the case of the interface being composed from half n and half p interfaces. For the cancellation to be effective, though, the width of the terraces should be smaller than the film thickness.

6. Summary

In this paper we have reviewed the present understanding on the origin of a 2DEG observed at the epitaxial interface between certain perovskite insulator thin films and certain perovskite insulator substrates, referring mostly to the paradigmatic case of an LAO film on STO. The ‘polar catastrophe’ concept and existing evidence for and against it have been reviewed. The polarization discontinuity of half a quantum of polarization (for LAO on STO) is obtained from the surface theorem of [43], and the symmetry of the bulk systems. This construction is related to other symmetry-protected topological phases appearing in other contexts, namely, the topological insulators related to other symmetries such as time-reversal and/or particle–hole symmetry. The discussion is generalized from the conventional (001) orientation of the interface to any orientation via the lattice of meaningful polarization vectors for these insulators, including vicinal (stepped) interfaces of high index. Alternative descriptions and formulations for the polar discontinuity found in the literature are reviewed and compared with the one based on formal polarization.

The polar catastrophe makes the pristine film unstable towards electric field screening mechanisms based on surface
and interface charges, the latter being the electron carriers of the 2DEG. The review of such instabilities is upheld by a simple model, following on from [39, 41], starting from the electronic reconstruction, followed by the redox defect model, and their possible coexistence. The model is generalized to arbitrary homogeneous surface–interface charging mechanisms.

Finally, the effects of deviations from the ideal epitaxial growth (interface effects) are reviewed, distinguishing between deviations respecting stoichiometry (such as inter-diffusion across the interface), which do not alter the qualitative situation described by the polar catastrophe and surface–interface charge screening, and deviations that do alter stoichiometry on a substantial scale.

Acknowledgments

We wish to thank H Y Hwang, S A T Redfern, P Ordejón, P Aguado, I Souza, R Resta, D Fontaine and D Vanderbilt for useful discussions. We acknowledge support from EPSRC and computing resources from CAMGRID in Cambridge and the Spanish Supercomputer Network (RES). This work was partly funded by UK’s EPSRC and the ARC project TheMoTherm (Grant No. 10/15-03). Work at Argonne was supported by DOE-DES under Contract No. DE-AC02-06CH11357. PhIG acknowledges a Research Professorship from the Francqui Foundation (Belgium).

References

[1] Ohtomo A and Hwang H Y 2004 *Nature* **427** 423–6
[2] Brinkman A, Huijben M, Van Zalk M, Huijben J, Zeitler U, Maan J C, Van der Wiel W G, Rijnders G, Blank D H A and Hilgenkamp H 2007 *Nature Mater.* **6** 493–6
[3] Reyren N et al 2007 *Science* **317** 1196–9
[4] Bert J, Kalisky B, Bell C, Kim M, Hikita Y, Hwang H and Moler K 2011 *Nature Phys.* **7** 676–71
[5] Li L, Richter C, Mannhart J and Ashoori R 2011 *Nature Phys.* **7** 676
[6] Basletic M, Maurice J L, Carretero C, Herranz G, Copie O, Ribes M, Jacquet E, Bouzehouane K, Fusil S and Barthelemy A 2008 *Nature Mater.* **7** 621–5
[7] Thiel S, Hammerl G, Schmehl A, Schneider C W and Mannhart J 2006 *Science* **313** 1942–5
[8] Chen C, Thiel S, Hammerl G, Schneider C W, Andersen K E, Heiberg C S, Mannhart J and Levy J 2008 *Nature Mater.* **7** 298–302
[9] Xiong H, Hikita Y, Bell C and Hwang H 2011 *Nature Commun.* **2** 494
[10] Irvin P, Ma Y, Bogorin D, Cen C, Bark C, Folkmann C, Eom C and Levy J 2010 *Nature Photon.* **4** 484–52
[11] Pallecchi I, Coda M, d’Agliano E G, Marré D, Caviglia A, Reyren N, Gariglio S and Triscone J M 2010 *Phys. Rev. B* **81** 085414
[12] Filippetti A, Delugas P, Verstraete M, Pallecchi I, Gadaleta A, Marré D, Li D, Gariglio S and Fiorentini V 2012 *Phys. Rev. B* **86** 195301
[13] Assmann E, Blaha P, Laskowski R, Held K, Okamoto S and Sangiovanni V 2013 *Phys. Rev. Lett.* **110** 078701
[14] Liang H et al 2013 *Sci. Rep.* **3** 1975
[15] Hwang H Y 2006 *Science* **313** 1895–6
[16] Pauli S A and Willmott P R 2008 *J. Phys.: Condens. Matter* **20** 264012
[17] Huijben M, Brinkman A, Koster G, Rijnders G, Hilgenkamp H and Blank D 2009 *Adv. Mater.* **21** 1665–77
[18] Mannhart J and Schlom D 2010 *Science* **327** 1607
[19] Pentecheva R and Pickett W 2010 *J. Phys.: Condens. Matter* **22** 043001
[20] Chen H, Kolpak A M and Ismail-Beigi S 2010 *Adv. Mater.* **22** 2881–99
[21] Schlom D and Mannhart J 2011 *Nature Mater.* **10** 168–9
[22] Chambers S A 2011 *Surf. Sci.* **605** 1133–40
[23] Pentecheva R, Arras R, Otte K, Ruiz V and Pickett W 2012 *Phil. Trans. R. Soc. A* **370** 4904–26
[24] Gabay M, Gariglio S, Triscone J M and Santander-Syro A 2013 *Eur. Phys. J. Spec. Top.* **222** 1177–83
[25] Nakagawa N, Hwang H and Muller D 2006 *Nature Mater.* **5** 204–9
[26] Janotti A, Bjaalie L, Gordon L and Van de Walle C 2012 *Phys. Rev. B* **86** 241108
[27] Reinle-Schmitt M et al 2012 *Nature Commun.* **3** 932
[28] Sing M et al 2009 *Phys. Rev. Lett.* **102** 176805
[29] Berner G et al 2010 *Phys. Rev. B* **82** 244105
[30] Takizawa M, Tsuda S, Susaki T, Hwang H Y and Fujimori A 2011 *Phys. Rev. B* **84** 245124
[31] Rubano A et al 2011 *Phys. Rev. B* **83** 155405
[32] Slooken E et al 2013 *Phys. Rev. B* **87** 085128
[33] Berner G et al 2013 *Phys. Rev. Lett.* **110** 247601
[34] Plum N et al 2013 arXiv:1304.5948
[35] Segal Y, Ngai J, Reiner J, Walker F and Ahn C 2009 *Phys. Rev. B* **80** 241107
[36] Berner G et al 2013 *Phys. Rev. B* **88** 115111
[37] Herranz G et al 2007 *Phys. Rev. Lett.* **98** 216803
[38] Kalabukhov A, Boikov Y, Serenkov I, Sakharov V, Börjesson J, Ljustina N, Olsson E, Winkler D and Claeson T 2011 *Europhys. Lett.* **93** 37001
[39] Bristowe N C, Littlewood P B and Artacho E 2011 *J. Phys.: Condens. Matter* **23** 081001
[40] Stengel M 2011 *Phys. Rev. B* **84** 205432
[41] Bristowe N, Artacho E and Littlewood P 2009 *Phys. Rev. B* **80** 045425
[42] Stengel M and Vanderbilt D 2009 *Phys. Rev. B* **80** 241103
[43] Vanderbilt D and King-Smith R 1993 *Phys. Rev. B* **48** 4442–55
[44] Murray É and Vanderbilt D 2009 *Phys. Rev. B* **79** 100102
[45] Goniakowski J, Finocchi F and Nogueira C 2008 *Rep. Prog. Phys.* **71** 016501
[46] King-Smith R D and Vanderbilt D 1993 *Phys. Rev. B* **47** 1651–4
[47] Marzari N and Vanderbilt D 1997 *Phys. Rev. B* **56** 12847–65
[48] Qi X L 2012 *Science* **338** 1550–1
[49] Resta R 1994 *Rev. Mod. Phys.* **66** 899
[50] Berry M 1984 *Proc. R. Soc.* A **383** 286–303
[51] Pauli S A and Willmott P R 2008 *Phys. Rev. B* **83** 045128
[52] Jadaun P, Xiao D, Niu Q and Banerjee S K 2013 *Phys. Rev. Lett.* **110** 085128
[53] Pentcheva R, Arras R, Otte K, Ruiz V and Pickett W 2012 *Phil. Trans. R. Soc. A* **370** 4904–26
[54] Aoki S, Tedy H, Ishikawa H and Ohtomo A 2009 *Phys. Rev. Lett.* **102** 078701
[55] Annadi A et al 2013 *Nature Commun.* **4** 1838
[121] Wang R, Fong D, Jiang F, Highland M, Fuoss P, Thompson C, Kolpak A, Eastman J, Streiffer S and Rappe A 2009 Phys. Rev. Lett. 102 47601
[122] Fong D et al 2006 Phys. Rev. Lett. 96 127601
[123] Spanier J, Kolpak A, Urban J, Grinberg I, Ouyang L, Yun W, Rappe A and Park H 2006 Nano Lett. 6 735–9
[124] Stephenson G B and Highland M J 2011 Phys. Rev. B 84 064107
[125] Bristowe N, Stengel M, Littlewood P, Pruneda J and Artacho E 2012 Phys. Rev. B 85 024106
[126] Stengel M 2011 Phys. Rev. Lett. 106 136803
[127] Seo H and Demkov A A 2011 Phys. Rev. B 84 045440
[128] Gu M, Wang J, Wu X and Zhang G 2012 J. Phys. Chem. C 116 24993–8
[129] Meevasana W, King P, He R, Mo S, Hashimoto M, Tamai A, Songsiriritthigul P, Baumberger F and Shen Z 2011 Nature Mater. 10 114–8
[130] Santander-Syro A et al 2011 Nature 469 189–93
[131] Chen Y, Pryds N, Kleibeuker J E, Koster G, Sun J, Stamate E, Shen B, Rijnders G and Linderoth S 2011 Nano Lett. 11 3774–8
[132] Lee S W, Liu Y, Heo J and Gordon R G 2012 Nano Lett. 12 4775–83
[133] Christensen D, Trier F, Chen Y, Smith A, Nygard J and Pryds N 2013 Appl. Phys. Lett. 102 021602
[134] Liu Z et al 2013 Phys. Rev. X 3 021010
[135] Chen Y et al 2013 Nature Commun. 4 1371