

# A new constituent of electrostatic energy in semiconductors

## An attempt to reformulate electrostatic energy in matter

Jean-Michel Sallese<sup>a</sup>

Swiss Federal Institute of Technology, Lausanne, Switzerland

Received 30 October 2015 / Received in final form 29 January 2016

Published online 1 June 2016

© The Author(s) 2016. This article is published with open access at [Springerlink.com](http://Springerlink.com)

**Abstract.** The concept of electric energy is revisited in detail for semiconductors. We come to the conclusion that the main relationship used to calculate the energy related to the penetration of the electric field in semiconductors is missing a fundamental term. For instance, spatial derivative of the electrostatic energy using the traditional formula fails at giving the correct electrostatic force between semiconductor based capacitor plates, and reveals unambiguously the existence of an extra contribution to the standard electrostatic free energy. The additional term is found to be related to the generation of space charge regions which are predicted when combining electrostatics with semiconductor physics laws, such as for accumulation and inversion layers. On the contrary, no such energy is needed when relying on electrostatics only, as for instance when adopting the so-called full depletion approximation. The same holds for neutral and charged insulators that are still consistent with the customary definition, but these two examples are in fact singular cases. In semiconductors for instance, this additional energy can largely exceed the energy gained by the dipoles, thus becoming the dominant term. This unexpected result clearly asks for a generalization of electrostatic energy in matter in order to reconcile basic concepts of electrostatic energy in the framework of classical physics.

## 1 Introduction

Interpretation of electric energy in conjunction with thermodynamics has been widely investigated, with a special interest for dielectric bodies and ideal conductors [1–5]. The electric energy stored inside of a body can be expressed whether in terms of charges and potentials restricted to the volume of the body [5], or in terms of fields including contributions beyond the physical boundary of the system, see relations (1) and (2) for linear polarisable systems [5]. In addition, careful considerations and exhaustive criticisms about the validity of electrostatic energy formulation in conductors and insulators has been addressed in references [1,2]. Beyond the thorough literature, instructive and complete analysis of thermodynamics of electric and magnetic fields has also been developed in references [1,2,4,6].

While the topic of electric energy in matter seems to be well established and widely accepted, we will bring some evidence that this paradigm must be revisited in semiconductors to say the least, resulting in a new contribution to the well-known Helmholtz free energy arising from the influence of the electric field.

Before discussing an instructive virtual experiment, we revisit some fundamental relationships dealing with elec-

trostatic energy from a general thermodynamic point of view.

The electrical work  $U_E$  that must be spent to gather charges from infinity into a volume  $\Omega$  is given by [3,5,6]:

$$U_E = \frac{1}{2} \int_{\Omega} \phi \rho d\Omega, \quad (1)$$

where  $\rho$  is the local charge density and  $\phi$  is the electric potential (note that the potential  $\phi$  must vanish at infinity [5]). The integral is limited to the volume of the body containing the charges, and in this sense, relation (1) represents the electric energy of the content of  $\Omega$  (assuming the system linear). The internal energy of electrical nature is then expected to be implicitly contained in (1), which is indeed how Frankl [7] analyzed the free energy stored in the depletion region of a silicon layer.

This formulation attributes energy to electric charges. Alternatively, adopting the electric field and displacement vectors concepts ( $\mathbf{E} = -\text{grad}(\phi)$ ,  $\text{div}(\mathbf{D}) = \rho$ , bold letters hold for vectors), the electrostatic energy can also be expressed from the electric field generated by the charges enclosed in the volume  $\Omega$ , provided that the integration is performed over the whole space  $\Omega_{\infty}$ , including matter:

$$U_E = \frac{1}{2} \int_{\Omega_{\infty}} \mathbf{E} \cdot \mathbf{D} d\Omega. \quad (2)$$

<sup>a</sup> e-mail: [jean-michel.sallese@epfl.ch](mailto:jean-michel.sallese@epfl.ch)

Again, the electric field and displacement vectors in (2) do not account for all sources in the universe; they are only assigned to the charges located in  $\Omega$ . It can be shown that equations (1) and (2) are equivalent and represent the electrostatic energy of those charges [5].

A generalisation of relation (2) to nonlinear polarisable materials [1–5] is also proposed:

$$U_E = \int_{\Omega_\infty} \left( \int_0^D \mathbf{E} \cdot d\mathbf{D} \right) d\Omega. \quad (3)$$

Relation (3) is more general as it gives the incremental work spent upon creating the electric field in matter and in free space, but without any assumption on the relationship between  $\mathbf{E}$  and  $\mathbf{D}$ , i.e. without assuming that the medium is linear. From a thermodynamic point of view, relations (2) and (3) represent the Helmholtz free electric energy of the system when assuming an isothermal process under constant deformation [1–3].

For an isotropic medium [5], the energy belonging to the dipoles [1,2] is also included in (3) through the polarization vector  $\mathbf{P}$  that satisfies  $\mathbf{D} = \epsilon_0 \cdot \mathbf{E} + \mathbf{P}$ . Finally, the electrostatic energy is made up of two contributions, i.e.  $U_E = U_f + U_P$ , where  $U_f$  is the electric field energy, valid in matter as in free space [2]:

$$U_f = \frac{1}{2} \epsilon_0 \int_{\Omega_\infty} |\mathbf{E}|^2 \cdot d\Omega \quad (4)$$

whereas  $U_P$  is an energy related to polarization processes experienced by the body  $\Omega$  only:

$$U_P = \int_{\Omega} \left( \int_0^{\mathbf{P}} \mathbf{E} \cdot d\mathbf{P} \right) d\Omega. \quad (5)$$

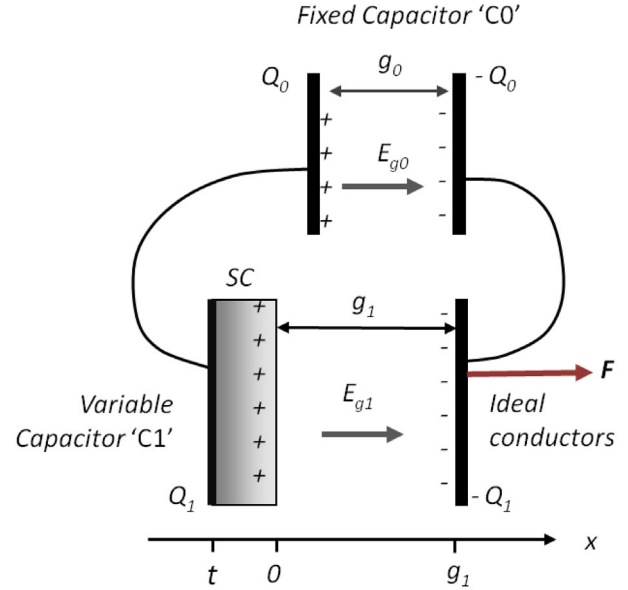
In dielectrics,  $U_P$  can be thought as a transformation of electric energy in some internal energy that belongs to the body. Obviously, this term cancels in ideal conductors since no electric field penetrates inside.

However, concerning semiconductors, we can wonder if  $U_P$  is still the only contribution to the Helmholtz free energy of electric nature.

In this work, we propose to analyze how the electric energy is transferred to semiconductors, and if this still follows the same law as for dielectrics. To the best of our knowledge, a detailed transfer of electric energy in semiconductors has never been examined so far.

## 2 Virtual experiment with semiconductor based capacitors

This section analyzes the work spent upon moving a semiconductor based capacitor plate with respect to a counter conductor plate, and compares it with the variation of the Helmholtz free energy as predicted from relations (1)



**Fig. 1.** Representation of the system of coupled capacitors used in the virtual experiment.  $\mathbf{F}$  represents the external force which is applied to the variable capacitor plate.

to (3). To this purpose, we assume two semi-infinite capacitors,  $C_0$  and  $C_1$ , with their plates connected such as in Figure 1. Except for  $C_1$  where one electrode is a semiconductor (p-type doped, without loss of generality), others electrodes are ideal conductors. This special arrangement makes the thermodynamic analysis self consistent as there is no need for introducing any external voltage source. The case of capacitors biased by means of an external voltage source has been inspected in many details by Bobbio [1,2].

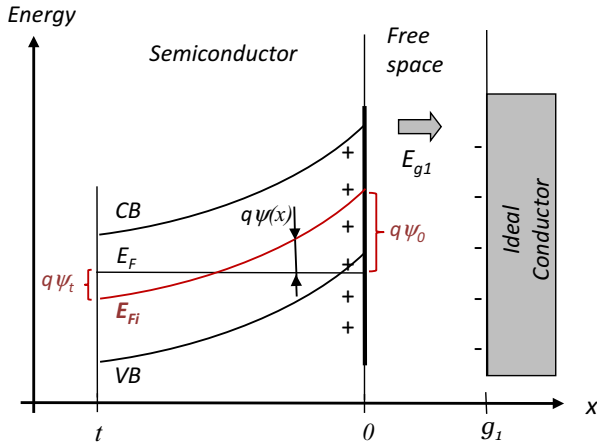
### 2.1 Regular derivation of the total electrostatic energy

Applying the definition of the electric energy given by relation (3) and assuming that electric charges on counter electrodes must compensate each other (we consider semi-infinite plates), we obtain the electric energy for each capacitor system  $C_0$  and  $C_1$  (scalars consistent with the axis orientation are used instead of vectors):

$$U_{C_0} = \int_{\Omega_0} \int_0^{D_0} E dD dx = \frac{\epsilon_0}{2} E_{g_0}^2 g_0 \quad (6)$$

$$U_{C_1} = \int_{\Omega_1} \int_0^{D_1} E dD dx = \frac{\epsilon_0}{2} E_{g_1}^2 g_1 + \int_t^0 \int_0^{D_1} E dD dx. \quad (7)$$

Here  $\Omega_0$  and  $\Omega_1$  represent the free space between the capacitor plates,  $g_0$  and  $g_1$  are the electrodes gaps for  $C_0$  and  $C_1$ ,  $E_{g_0}$  and  $E_{g_1}$  are the *uniform* (scalar) electric fields between the electrodes (according to Fig. 1  $E_{g_0}, E_{g_1} \geq 0$ ) and “ $t$ ” is the semiconductor thickness (note that  $t$  is negative given the origin of the axis). In addition, since we suppose that the polarisation of the semiconductor is linear, i.e.  $D = \epsilon_{SC} E$  ( $\epsilon_{SC}$  is the semiconductor dielectric



**Fig. 2.** Energy representation of the conductor-vacuum-semiconductor capacitor structure in a general case without assuming the presence of a neutral region ( $E_F$  and  $E_{Fi}$  are respectively the Fermi and intrinsic Fermi energies).

constant), relation (7) becomes (according to Fig. 2,  $E_{g1}$  is positive while  $\psi$  is negative):

$$\begin{aligned}
 U_{C1} &= \frac{\epsilon_0}{2} E_{g1}^2 g_1 + \frac{\epsilon_{SC}}{2} \int_t^0 E^2 dx \\
 &= \frac{\epsilon_0}{2} E_{g1}^2 g_1 - \frac{\epsilon_{SC}}{2} \int_t^0 E \frac{d\psi}{dx} dx \\
 &= \frac{\epsilon_0}{2} E_{g1}^2 g_1 - \frac{\epsilon_{SC}}{2} \int_{\psi_t}^{\psi_0} E d\psi, \tag{8}
 \end{aligned}$$

where  $\psi_0$  and  $\psi_t$  are the potentials evaluated at the boundaries of the semiconductor body such as shown in Figure 2, which implicitly assumes the Fermi potential as reference. We will come to this point later. Note that the last equality in relation (8) holds if the electric field takes a unique value for a given potential. However, this could be unfulfilled when the potential exhibits extrema across the semiconductor layer. In that case, the integral should be split in different domains wherein a unique correspondence exists between  $E$  and  $\psi$ .

Then, the total electrostatic energy  $U_E$  of the capacitors system is readily obtained as a function of the electric fields and potentials in the semiconductor:

$$\begin{aligned}
 U_E &= U_{C0} + U_{C1} \\
 &= \frac{\epsilon_0}{2} E_{g0}^2 g_0 + \frac{\epsilon_0}{2} E_{g1}^2 g_1 - \frac{\epsilon_{SC}}{2} \int_{\psi_t}^{\psi_0} E d\psi. \tag{9}
 \end{aligned}$$

Before proceeding further, a discussion to clarify the meaning of the potential  $\psi$  is needed.

When expressing the electric field as the gradient of a potential, the origin for that potential is a priori not relevant. However, when translating the integral over space

into an integral over the potential as in relation (8), the potential  $\psi$  should be defined in such a way that it remains consistent with the expression of the electric field. A typical potential profile is shown in Figure 2 where the origin for potentials is taken at the constant Fermi potential (here  $\psi_0 < 0$  and  $\psi_t > 0$ ). But this choice could be different and only after the link between the electric field and the potential has been established, the origin for the potential peculiar to relations (8) and (9) can be set.

After presenting a quite general treatment, the special case of a uniformly doped semiconductor in which a neutral region is recovered when moving away from the free surface will be analyzed in details.

## 2.2 Electric work upon electrode displacement: need for a new energy term

In this virtual experiment, before being isolated, tied electrodes are pre-charged with a total average charge density (per unit surface)  $Q_T = Q_0 + Q_1$  (see Fig. 1). Any displacement  $d\mathbf{x}$  of the semiconductor counter electrode (others are supposed fixed) will induce a variation of the electrostatic energy for  $C0$  and  $C1$ ; as well as a mechanical work  $\delta W_F$  arising from the attractive electrostatic force  $\mathbf{F}$  that exists between the plates.

Invoking the fundamental law of thermodynamics, when the displacement is performed at constant total charge, i.e. no connection to any voltage source, we can write:

$$\begin{aligned}
 \delta W_F &= -\mathbf{F} \cdot d\mathbf{x} \\
 &= dU_f + \sum_{conductors} dU_C + \sum_{semiconductors} dU_{SC} \\
 &\quad - \sum_{conductors} T_C dS_C - T_{SC} dS_{SC}. \tag{10}
 \end{aligned}$$

As in reference [2], we introduce the Helmholtz free energy for each of the bodies. Relation (10) becomes:

$$\begin{aligned}
 \delta W_F &= dU_f + \sum_{conductors} dA_C + \sum_{conductors} S_C \cdot dT_C \\
 &\quad + dA_{SC} + S_{SC} \cdot dT_{SC}, \tag{11}
 \end{aligned}$$

where  $dU_f$  is the variation of the electric field energy in the whole space (including bodies), while  $dA_C$  and  $dA_{SC}$  are the variations of the Helmholtz free energy of the conductor and semiconductor plates induced by the displacement  $d\mathbf{x}$  while maintaining constant the total charge  $Q_T$ . Finally,  $S_C$  and  $S_{SC}$  represent the entropy of each subsystem. Note that for dielectrics, the Helmholtz free energy reverts to  $U_P$  [2].

Still as in reference [2], we assume that the temperature of the system is maintained fixed ( $T_C$  and  $T_{SC}$  are equal and constant). Assuming that there is no deformation, and since the temperature is constant, no variation

of the Helmholtz free energy can arise in conducting bodies [2], i.e.  $dA_C = 0$ . Then, relation (11) simplifies into:

$$\delta W_F = dU_f + dA_{SC}. \quad (12)$$

Bobbio [2] demonstrated that the term in the right hand side of (12) is the electric energy as defined from relation (3):

$$\delta W_F = dU_E. \quad (13)$$

The link between the mechanical work and energies of electrical nature is now in order.

The way  $Q_T$  will redistribute among the connected electrodes will depend on the total charge density itself and on the electrodes gaps, i.e.  $g_0$  and  $g_1$ ; as well as on the physical nature of the plates, i.e. conductors or semiconductors.

As discussed before, we start considering a very general situation where the semiconductor plate is not necessarily neutral at  $x = t$  (in this case, there could also be a charge sheet layer at the contact-semiconductor interface at  $x = t$  since the electric field must vanish for  $x < t$ ). To be consistent with further developments, we will call  $\psi_0$  the *surface potential* and use the notation  $\psi_s$  instead.

Given that connected plates share the same potential, we have ( $E_{g0}, E_{g1} \geq 0$  and  $\psi_S - \psi_t \leq 0$ ) (see Fig. 2):

$$E_{g0}g_0 = E_{g1}g_1 - (\psi_S - \psi_t). \quad (14)$$

Since the sum of the charge densities on each pair of tied electrodes is fixed, we call it  $Q_T$ , Gauss theorem imposes that the sum of the electric fields  $E_{g1}$  and  $E_{g0}$  is invariant:

$$\delta(E_{g0} + E_{g1}) = 0. \quad (15)$$

Next, from the continuity of the displacement vector, without presuming for any fixed charge sheet layer on the semiconductor surface ( $x = 0$ ), the electric fields across the semiconductor/free space interface satisfy:

$$\varepsilon_{SC}E_s = \varepsilon_0E_{g1}, \quad (16)$$

where  $E_S$  is the surface electric field evaluated *inside* the semiconductor at  $x = 0$ .

Differentiating (9) gives the variation of the electric energy for the capacitors system:

$$\begin{aligned} dU_E &= (\varepsilon_0E_{g0}dE_{g0}g_0) \\ &+ \left( \varepsilon_0E_{g1}dE_{g1}g_1 + \frac{\varepsilon_0}{2}E_{g1}^2dg_1 \right) \\ &- \frac{\varepsilon_{SC}}{2}d \int_{\psi_t}^{\psi_s} E d\psi. \end{aligned} \quad (17)$$

Noting that the electric field takes a unique value for a given potential, the last term in (17) simplifies:

$$\frac{\varepsilon_{SC}}{2}d \int_{\psi_t}^{\psi_s} E d\psi = \frac{\varepsilon_{SC}}{2}(E_S d\psi_S - E_t d\psi_t), \quad (18)$$

where  $E_S$  and  $E_t$  are the electric fields at  $x = 0$  and  $x = t$ , still inside the semiconductor.

Next, merging relations (14), (15) and (16) with (17), the change in the electric energy can be expressed in terms of  $C1$  capacitor quantities only:

$$\begin{aligned} dU_E &= \varepsilon_{SC}dE_S(\psi_S - \psi_t) - \frac{\varepsilon_{SC}}{2}(E_S d\psi_S - E_t d\psi_t) \\ &+ \left( \frac{\varepsilon_0}{2}E_{g1}^2dg_1 \right). \end{aligned} \quad (19)$$

Finally, the force per unit area acting on the semiconductor plate as derived from the electric energy is:

$$\begin{aligned} F_U &= \frac{dU_E}{dg_1} = \varepsilon_{SC} \frac{dE_S}{dg_1} (\psi_S - \psi_t) \\ &- \frac{\varepsilon_{SC}}{2} \left( E_S \frac{d\psi_s}{dg_1} - E_t \frac{d\psi_t}{dg_1} \right) + \left( \frac{\varepsilon_0}{2}E_{g1}^2 \right). \end{aligned} \quad (20)$$

This is what we call the ‘*field energy force*’ since it is derived from the electric field and dipole energies that are already included in relation (3).

On the other hand, the electric charge density on  $C1$  creates an attractive force between the plates through the Coulomb force  $F_C$ . This force per unit area is the product of the total charge density in the semiconductor times the electric field generated by the counter electrode, which is half the electric field in the gap since we must exclude the contribution to the field induced by the charge itself ( $F_C$  is positive according to Fig. 1):

$$F_C = (\varepsilon_0E_{g1}) \frac{1}{2}E_{g1} = \frac{1}{2}\varepsilon_0E_{g1}^2. \quad (21)$$

Relation (21) defines the ‘genuine’ electrostatic force [5]; as such it should be regarded as the actual force. We expect that the force derived from the free energy (relation (20)) and the Coulomb force should be strictly equal. However, relations (20) and (21) are actually not equivalent.

*The unexpected inconsistency between these two formula is a major result per se and suggests that the definition of the electric energy as given by relation (3) does not represent the total electrostatic energy gained by the system.*

Without loss of generality, a new contribution is introduced in the Helmholtz free energy. We define this extra quantity by  $A_{Extra}$ . Then, the new electric energy  $U_E^{new}$  for the system of coupled capacitors writes:

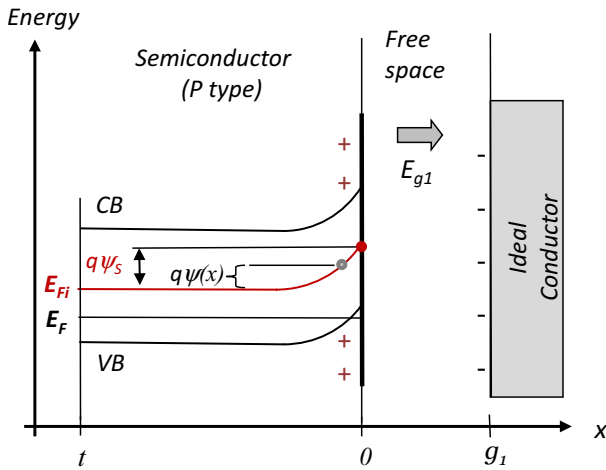
$$U_E^{new} = U_E + A_{Extra} \quad (22)$$

and relation (13) becomes:

$$\delta W_F = dU_E^{new} = dU_f + dA_{SC} + dA_{Extra}. \quad (23)$$

Now, if we impose  $F_C$  and  $F_U$  to be equal, the extra energy must satisfy the following differential equation:

$$\begin{aligned} \frac{dA_{Extra}}{dg_1} + \varepsilon_{SC} \frac{dE_S}{dg_1} (\psi_S - \psi_t) \\ - \frac{\varepsilon_{SC}}{2} \left( E_S \frac{d\psi_s}{dg_1} - E_t \frac{d\psi_t}{dg_1} \right) = 0. \end{aligned} \quad (24)$$



**Fig. 3.** Energy representation of the conductor-vacuum-semiconductor capacitor structure when a neutral region is recovered away from the free surface. The surface potential is the potential drop inside the semiconductor ( $E_F$  and  $E_{Fi}$  are respectively the Fermi and intrinsic Fermi energies).

Integrating over  $g_1$  from  $-\infty$  to some finite value, and noting that at infinity the electric field  $E_S$  must vanish since relation (14) cannot diverge, the extra energy is given by:

$$A_{Extra} = \frac{\epsilon_{sc}}{2} \cdot \int_{\psi_S^\infty}^{\psi_S} E_S \cdot d\psi - \frac{\epsilon_{sc}}{2} \int_{\psi_t^\infty}^{\psi_t} E_t \cdot d\psi - \epsilon_{sc} \cdot \int_0^{E_S} (\psi_S - \psi_t) \cdot dE, \quad (25)$$

where  $\psi_S^\infty$  and  $\psi_t^\infty$  are the potentials evaluated at  $x = 0$  and  $x = t$  when the gap separation  $g_1$  goes to infinity. These can be calculated once the physical parameters of the semiconductor system are known. In addition, we implicitly assume that  $A_{Extra} = 0$  when  $g_1 \rightarrow \infty$ . This extra energy involves only physical quantities belonging to the semiconductor (electric fields, potentials and dielectric constant). Then, likewise for the energy of the dipoles  $U_P$ ,  $A_{Extra}$  is also part of the semiconductor free energy.

Relation (25) expresses the additional electric energy in a semiconductor layer. This result was not anticipated if we concede that relation (3) was held to account for the energy of electric nature in matter and in free space. In the next section, a special case where a neutral region is recovered inside the semiconductor layer is illustrated in details.

### 2.3 The case of a partially depleted semiconductor

Without loss of generality, we assume a p-type doped semiconductor layer with a doping density  $N_A$ . An energy representation is shown in Figure 3 (the semiconductor is set to a lower potential with respect to the counter electrode). In addition, we consider that the semiconductor is neutral at  $x = t$ , and therefore  $E_t = 0$  (note that for doping densities greater than  $10^{15} \text{ cm}^{-3}$ , neutrality is recovered after

some micrometers only). In this case, it is quite common to define the intrinsic Fermi potential in the neutral body as the origin for the potentials. Therefore we have  $\psi_t = 0$  and the extra energy simplifies:

$$A_{Extra} = \frac{\epsilon_{sc}}{2} \cdot \int_0^{\psi_S} E_S \cdot d\psi - \epsilon_{sc} \cdot \int_0^{E_S} \psi_S \cdot dE = \epsilon_{sc} \cdot \left( \frac{3}{2} \cdot \int_0^{\psi_S} E_S \cdot d\psi - E_S \cdot \psi_S \right). \quad (26)$$

Introducing the susceptibility  $\chi = \epsilon_{sc}/\epsilon_0 - 1$  and assuming the medium linear, the polarization vector can be written  $\mathbf{P} = \epsilon_0 \cdot \chi \cdot \mathbf{E}$ . From relation (5) the dipoles free energy is:

$$U_P = \frac{-\epsilon_0 \chi}{2} \int_0^{\psi_S} E_S d\psi. \quad (27)$$

Combining (26) with (27) gives:

$$A_{Extra} = -\epsilon_{sc} \left( \psi_S E_S + \frac{3}{\epsilon_0 \chi} U_P \right). \quad (28)$$

It is instructive to define the energy contribution restricted to the volume of the semiconductor,  $U_{SC}$  (note that this is not the internal energy, see [1] for clarification). According to relation (2), we have:

$$U_{SC} = \frac{1}{2} \cdot \int_{\Omega_{SC}} \mathbf{E} \cdot \mathbf{D} \cdot d\mathbf{x} = -\frac{\epsilon_{sc}}{2} \cdot \int_0^{\psi_S} \mathbf{E} \cdot d\psi. \quad (29)$$

In addition to the field and dipoles energies, the new energy  $U_{SC}^{new}$  still restricted to the semiconductor plate must now incorporate the extra energy  $A_{Extra}$ :

$$U_{SC}^{new} = A_{Extra} + U_{SC} = \frac{\epsilon_{sc}}{2} \int_0^{\psi_S} E_S d\psi - \epsilon_{sc} \int_0^{E_S} \psi_S dE - \frac{\epsilon_{sc}}{2} \int_0^{\psi_S} E d\psi. \quad (30)$$

Noting that the link between  $E$  and  $\psi$  is unique and independent of the coordinate, using  $E_S$  or  $E$  in relation (30) makes no difference for the integrals and relation (30) simplifies:

$$U_{SC}^{new} = -\epsilon_{sc} \int_0^{E_S} \psi_S dE. \quad (31)$$

As expected,  $U_{SC}$  and  $U_{SC}^{new}$  are not equivalent.

Adding to  $U_{SC}^{new}$  the contribution of the remaining electric energy stored in the free space surrounding the semiconductor ( $\Omega - \Omega_{SC}$ ), the total electrostatic energy for the system of capacitors  $U_E^{new}$  is readily obtained:

$$U_E^{new} = \frac{1}{2} \cdot \int_{\Omega - \Omega_{SC}} \mathbf{E} \cdot \mathbf{D} \cdot d\mathbf{x} - \epsilon_{sc} \cdot \int_0^{E_S} \psi_S \cdot dE. \quad (32)$$

This electric energy is different from what we used initially, see relation (9). We can gain more insight about the meaning of  $U_{SC}$  and  $U_{SC}^{new}$  when relying on the standard  $E$ - $\Psi$  plot used to model the Metal-Insulator-Semiconductor capacitors system [8] that links the electric field  $E$  to the potential  $\psi$  (assuming non degenerate semiconductors):

$$E(\psi) = -\text{sign}(\psi) \sqrt{\frac{2U_T q N_A}{\epsilon_{SC}}} \times \sqrt{\frac{n_i^2}{N_A^2} \left( e^{\frac{\psi}{U_T}} - \frac{\psi}{U_T} - 1 \right) + \left( e^{-\frac{\psi}{U_T}} + \frac{\psi}{U_T} - 1 \right)}, \quad (33)$$

where  $n_i$  is the intrinsic carrier density and  $U_T$  is the thermodynamic potential ( $U_T = k_B T/q$ ).

This representation is very adequate to explain the distinction between the ‘typical’ and the ‘new’ formulations of electric energy restricted to the semiconductor body,  $U_{SC}$  and  $U_{SC}^{new}$ .

Basically, the new expression stated by relation (31) represents the area (blue) above the curve in the  $E$ - $\Psi$  plot shown in Figure 4a, whereas the contribution given by the standard relation (29) is representing half of the lower area (green). When the surface potential is below the onset of strong inversion as shown in Figure 4a, the upper area is almost half of the bottom one, and it is likely that both definitions converge to the same value. However, when the surface potential is beyond the strong inversion limit as in Figure 4b, the upper area dominates, indicating that relations (29) and (31) are not equivalent.

### 3 The case of neutral and charged insulators

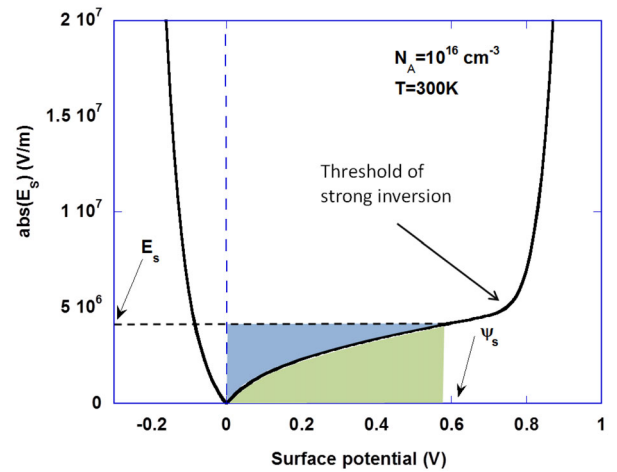
There are situations where the electrostatic force derived from the electrostatic energy given from relations (1) to (3) works correctly. We will provide some evidence that this does happen for insulators, and more generally for charged insulators, which are de facto traditional systems coming along with electric energy considerations.

Substituting the semiconductor for an insulator in the system of coupled capacitors sketched in Figure 1, we calculate the Coulomb  $F_C$  and energy based  $F_U$  forces and compare them. Regarding the energy stored in the capacitor  $C1$ , relation (7) can be rewritten in terms of the electric field in the insulator  $E_i$  ( $\epsilon_i$  is the insulator dielectric constant):

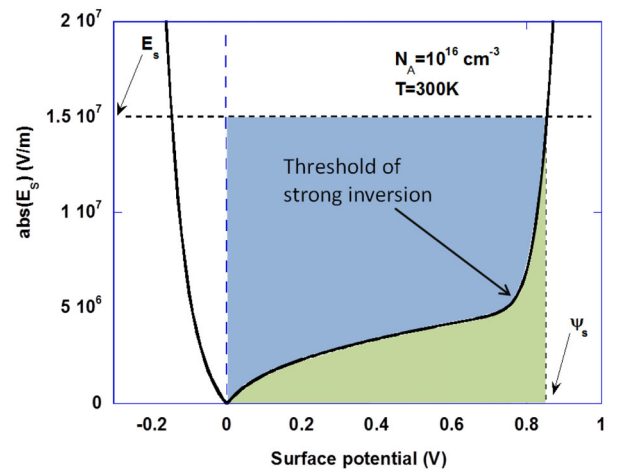
$$U_{C1} = \frac{\epsilon_i}{2} \int_t^0 E_i^2(x) dx + \frac{\epsilon_0}{2} E_{g1}^2 g_1. \quad (34)$$

Introducing the local charge density in the insulator  $\rho(x)$  and assuming that  $\rho(x)$  does not depend on the local potential, integrating Poisson equation in the insulator gives:

$$E_i(x) = \frac{1}{\epsilon_i} \int_0^x \rho(u) du + E_i(0). \quad (35)$$



(a)



(b)

**Fig. 4.** Electric field versus the surface potential for a p-type doped semiconductor layer. The upper (blue) and lower (green) areas delimited by the  $E$ - $\Psi$  curve represent the integrals defined in the first line of relation (26). Panel (a) illustrates the case when the semiconductor is in depletion. In this case, the upper area is approximately half of the lower one. However, when strong inversion takes place as in panel (b), the upper area dominates, illustrating the occurrence of the extra energy of electric nature.

Since the charge density depends on the coordinate only, we can express the electric field in the form:

$$E_i(x) = f(x) + E_i(0), \quad (36)$$

where  $f(x)$  is a function that is independent of the electric potential.

Similarly, integration of (36) gives the potential distribution in the insulator:

$$\psi(x) = \psi(0) - \int_0^x f(u) du - x E_i(0). \quad (37)$$

After rearrangement, the total electric energy  $U_E = U_{C1} + U_{C0}$  can be written as:

$$U_E = \frac{\varepsilon_i}{2} \left[ \int_t^0 f^2(u) du + 2E_i(0) \int_t^0 f(u) du - E_i^2(0)t \right] + \frac{\varepsilon_0}{2} E_{g1}^2 g_1 + \frac{\varepsilon_0}{2} E_{g0}^2 g_0. \quad (38)$$

Noting that integrals in the bracket of (38) do not depend on the local potential, and thus on the value of the gap  $g_1$ , the force arising from the derivative of the electric energy is:

$$F_U = \frac{dU_E}{dg_1} = \varepsilon_i \frac{dE_i(0)}{dg_1} \int_t^0 f(u) du - \varepsilon_i t \frac{dE_i(0)}{dg_1} E_i(0) + \frac{\varepsilon_0}{2} E_{g1}^2 + \varepsilon_0 E_{g1} \frac{dE_{g1}}{dg_1} g_1 + \varepsilon_0 E_{g0} \frac{dE_{g0}}{dg_1} g_0. \quad (39)$$

Additionally, connected plates must share the same potential, which from relation (37) gives:

$$E_0 g_0 = E_1 g_1 + \psi(t) - \psi(0) = E_1 g_1 - \int_0^t f(u) du - t E_i(0). \quad (40)$$

Introducing the identity (40) in relation (39), we obtain:

$$\frac{dU_E}{dg_1} = \varepsilon_i \frac{dE_i(0)}{dg_1} \int_t^0 f(u) du - \varepsilon_i \frac{dE_i(0)}{dg_1} E_i(0) t + \frac{\varepsilon_0}{2} E_{g1}^2 + \varepsilon_0 E_{g1} \frac{dE_{g1}}{dg_1} g_1 + \varepsilon_0 \frac{dE_{g0}}{dg_1} \left( E_1 g_1 - \int_0^t f(u) du - t E_i(0) \right). \quad (41)$$

Again, charge conservation writes:

$$dE_{g1} = -dE_{g0}. \quad (42)$$

Similarly, the continuity of the displacement vector at the dielectric-gap boundary gives:

$$\varepsilon_i E_i(0) = \varepsilon_0 E_{g1}. \quad (43)$$

Finally, using relations (42) and (43) in (41), we find that the electrostatic force based on the electric energy variation upon the electrode displacement reverts to the Coulomb force between the charged capacitor plates:

$$\frac{dU_E}{dg_1} = F_U = \frac{1}{2} \varepsilon_0 E_{g1}^2. \quad (44)$$

We can then conclude that as far as insulators are concerned, there is no need to introduce any new free energy of electric nature and the usual definition is leading to the correct result. This apparently 'trivial' finding for insulators could explain why the inconsistency pointed out in semiconductors has been concealed.

#### 4 A sufficient condition for the existence of the additional free energy

In the quest for a more general criterium, we can analyze when the extra energy given by (26) vanishes, i.e. when the electric energy of the capacitor system can still be obtained from relation (3). Following former analysis, this condition is verified as soon as  $F_U = F_C$  in relations (20) and (21). Imposing this identity links surface electric fields to potentials at front and back interfaces.

$$\varepsilon_{SC} \frac{dE_S}{dg_1} (\psi_S - \psi_t) - \frac{\varepsilon_{SC}}{2} \left( E_S \frac{d\psi_S}{dg_1} - E_t \frac{d\psi_t}{dg_1} \right) = 0. \quad (45)$$

Additionally, when the semiconductor layer is neutral at  $x = t$ , this condition simplifies further:

$$\frac{dE_S}{E_S} = \frac{1}{2} \frac{d\psi_S}{\psi_S}. \quad (46)$$

Since there is a one-to-one correspondence between the local electric field and the local potential, see relation (33) for instance, this condition can also be extended to any coordinate inside the body, leading to the differential equation:

$$\frac{dE(x)}{E(x)} = \frac{1}{2} \frac{d\psi(x)}{\psi(x)}. \quad (47)$$

Solutions of (47) are readily obtained:

$$E(x) = C \sqrt{-\psi(x)} \quad \text{when} \quad \psi(x) \leq 0 \quad \text{or} \quad (48a)$$

$$E(x) = -C \sqrt{\psi(x)} \quad \text{when} \quad \psi(x) \geq 0, \quad (48b)$$

where  $C$  is a positive valued integration constant.

Without loss of generality, we assume  $\psi_S \leq 0$  as in Figure 3 so that  $E(x) = C \sqrt{-\psi(x)}$ .

Making use of the Poisson equation, we find that the charge density must be constant (and negative in our case) in the semiconductor where the electric field is not null:

$$\frac{dE(x)}{dx} = \frac{-\rho(x)}{\varepsilon} = \frac{C^2}{2} \quad (49)$$

(when  $\psi_S \geq 0$  the same conclusion applies with a positive charge).

Basically, for semiconductors this condition reverts to the so-called full depletion approximation in a uniformly doped material [8]. It imposes that the body is fully depleted down to a given coordinate until it changes for neutrality in a step-like transition. In fact, this analysis makes use of the Poisson equation only, discarding the Fermi-Dirac or Boltzmann statistics, which is a quite common

approximation for semiconductors operating in depletion mode.

We can illustrate this result by considering that the semiconductor layer (doped  $N_A$ ) in Figure 1 gets fully depleted from the surface down to a coordinate  $x_0$ , beyond which it recovers neutrality, meaning that  $\psi(x_0) = 0$  and  $E(x_0) = 0$ . Solving the Poisson equation while assuming a charge density  $-qN_A$  ( $-q$  is the electron charge) between  $[x_0, 0]$  and 0 elsewhere, the potential and electric field at the surface are readily obtained:

$$\psi_S = -\frac{1}{2} \frac{qN_A}{\varepsilon_{SC}} x_0^2 \quad \text{and} \quad E_S = \frac{qN_A}{\varepsilon_{SC}} x_0. \quad (50)$$

Therefore the surface electric field and the surface potential are linked as follows:

$$E_S = \sqrt{\frac{2qN_A}{\varepsilon_{SC}}} \sqrt{-\psi_S}. \quad (51)$$

Evaluation of integrals in relation (26) gives:

$$\frac{\varepsilon_{SC}}{2} \int_0^{\psi_S} E_S d\psi = \frac{\varepsilon_{SC}}{3} \sqrt{\frac{2qN_A}{\varepsilon_{SC}}} (-\psi_S)^{3/2}, \quad (52)$$

$$\varepsilon_{SC} \int_0^{E_S} \psi_S dE = \frac{\varepsilon_{SC}}{3} \left( \frac{2qN_A}{\varepsilon_{SC}} \right)^{-1} E_S^3. \quad (53)$$

Apart for the sign, these quantities are equal, implying that the extra energy does vanishes in this case. However, as it will be illustrated in the next section, adopting a more consistent approach involving Fermi-Dirac or Boltzmann statistics, which rules out this quite crude full depletion approximation, will highlight a new energy term. The additional energy gained by the semiconductor is then likely to have its origin in statistical physics.

## 5 Some tangible quantities: energies and forces in an ideal semiconductor based capacitor

In order to estimate the magnitude and impact of the free energy given by relation (26), we rely on the analytical expression that links the surface potential to the surface electric field in a semiconductor plate. Considering a non-degenerate p-type doped silicon, the surface potential and the surface electric field satisfy the following well-known relationship, valid for depletion, inversion and accumulation [8] (signs are consistent with Fig. 3):

$$E_S = -\text{sign}(\psi_S) \sqrt{\frac{2U_T q N_A}{\varepsilon_{SC}}} \times \sqrt{\frac{n_i^2}{N_A^2} \left( e^{\frac{\psi_S}{U_T}} - \frac{\psi_S}{U_T} - 1 \right) + \left( e^{-\frac{\psi_S}{U_T}} + \frac{\psi_S}{U_T} - 1 \right)}, \quad (54)$$

where  $U_T$  is the thermal voltage,  $n_i$  is the intrinsic carrier density and  $N_A$  is the doping concentration, other symbols having their usual meaning.

### 5.1 Computation of the extra energy in non-degenerate silicon

As already discussed, adopting the conventional representation  $E_S(\psi_S)$  and considering only depletion-inversion modes (positive values of the surface potential in this case), relation (26) translates into a simple picture:  $A_{Extra}$  is merely half of the lower area (green) minus the upper one (blue) (see Figs. 4a and 4b). This graphical interpretation of  $A_{Extra}$  is useful to appraise the meaning of the ‘extra energy’. For instance, according to relations (52) and (53), the area in the upper side of the curve (blue) is about half the area below (green) as long as the surface potential  $\Psi_S$  remains below 0.7 volt (see Fig. 4a), which is the threshold of strong inversion in our case. This is where the potential and the electric field satisfy the full depletion approximation to a good extent. Indeed, while  $\Psi_S$  remains below 0.7 volt, holes can be neglected while keeping only the linear term in  $\psi_s$  under the square root in relation (54), leading to:

$$E_S = -\sqrt{2qN_A/\varepsilon_{SC}} \sqrt{\psi_S}, \quad (55)$$

which is formally the same as relation (48b).

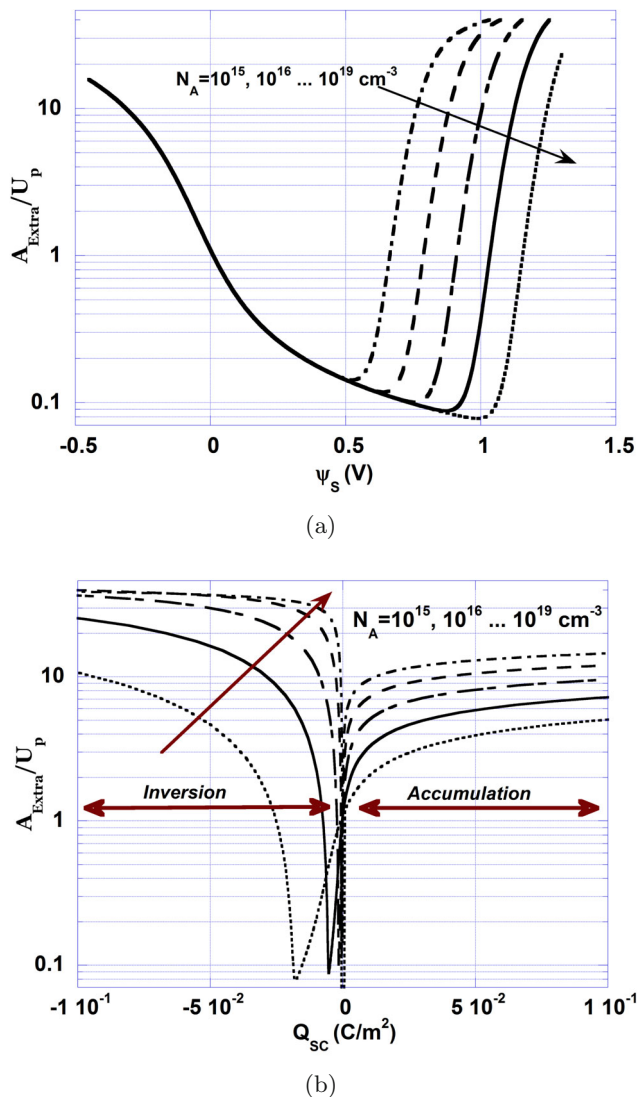
However, increasing the surface potential beyond that limit (0.7 volt) increases the weight of the blue area with respect to the green one (see Fig. 4b), and predicts that the two definitions of electric energy given by relations (29) and (31) will start to deviate from one another. In fact the amount of the additional free energy  $A_{Extra}$  will increase exponentially with the surface potential.

Similarly, for negative values of  $\Psi_S$ , i.e. accumulation mode, we can show that the additional free energy  $A_{Extra}$  is never negligible whatever the value of the surface potential (the surface ‘above’ the curve is always greater than half of the ‘bottom’ counterpart). Therefore, we anticipate that as soon as the surface potential matches with inversion or accumulation,  $A_{Extra}$  will never be negligible. All the more, it will be the dominant contribution to the total electrostatic free energy.

To perceive how the Helmholtz free electric energy is shared between the two ‘components’  $A_{Extra}$  and  $U_P$ , we evaluate the ratio  $A_{Extra}/U_P$  versus the surface potential for various doping densities (see Fig. 5a). The additional electric energy upon a change in the surface potential dominates when the semiconductor is set in accumulation, i.e.  $\psi_S < 0$  V, or above the threshold of inversion (i.e.  $\psi_S^{th} > 0.7$  V for  $N_A = 10^{16}$  cm<sup>-3</sup>). In inversion mode, this ‘threshold’  $\psi_S^{th}$  is shifted towards higher values when the doping density is increased, whereas in accumulation the doping density has almost no effect when illustrated with respect to the surface potential. Note that before reaching the strong inversion limit (which depends on the doping density), the ratio  $A_{Extra}/U_P$  remains lower than unity, meaning that the extra energy is minimized, which is expected since this is when the full depletion approximation is satisfied to a good extent.

A somehow different representation of  $A_{Extra}/U_P$  in terms of the charge density stored in the semiconductor is shown in Figure 5b. Except in a limited region where





**Fig. 5.** Ratio between the extra free energy  $A_{Extra}$  and the dipoles energy  $U_P$  versus the surface potential (a) and versus the charge density (b) in a silicon layer, for different doping concentrations. In (b), we note that in accumulation,  $A_{Extra}$  is always dominant, in contrast to depletion-inversion mode where it can have a minor contribution. At flat band, the ratio takes a well defined value independent of the doping density.

$-Q_{sc}$  is lower than about  $10^{-3} \text{ C/cm}^2$  (depletion), for a given space charge density the energy ratio increases with the doping level. In addition,  $A_{Extra}$  exceeds always  $U_P$  in accumulation, which is not the case in depletion-inversion mode where it can be relatively small, but still never cancels. Interestingly, when the surface potential is close to the flat band ( $\psi_s \approx 0 \text{ V}$ ), the extra energy tends to an asymptotic value independent of the doping density,  $A_{Extra}/U_P \approx 3\varepsilon_{sc}/\varepsilon_0 \cdot \chi$  (about 1.09 for silicon). Then, the additional electrostatic energy stored in a quasi neutral silicon layer still represents about half of the electric field and dipoles energies restricted to that silicon body.

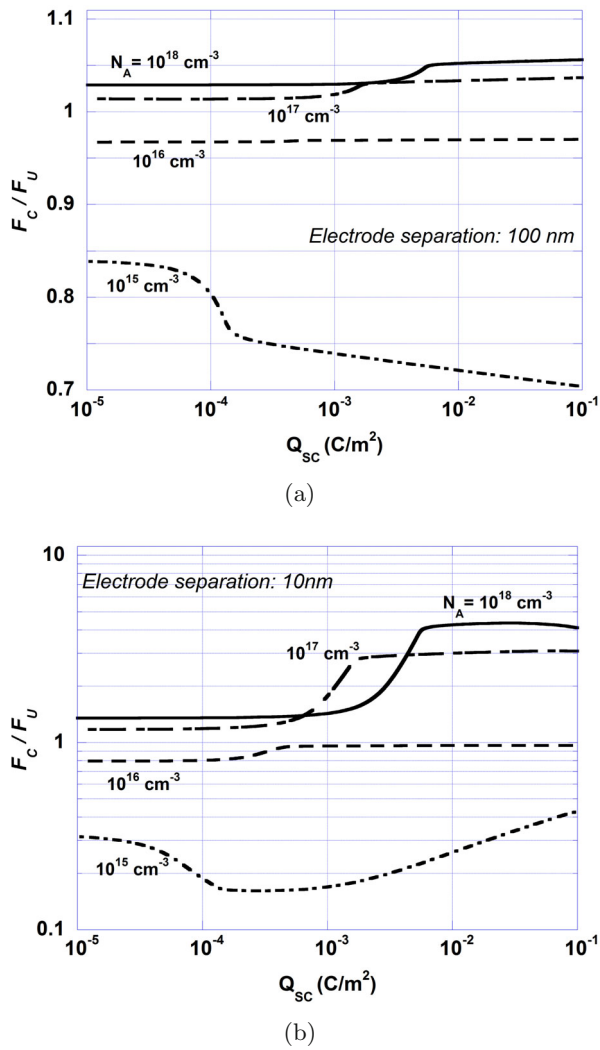
## 5.2 Appraisal of Coulomb and energy based electrostatic forces

The aim of this section is to illustrate the extra energy contribution through the tangible quantity of electrostatic force evaluated in a nanometer scale capacitor systems.

To that purpose, two values for the electrode gap of  $C1$  are chosen, namely 100 and 10 nm, as well as different doping densities. Without loss of generality, for each case we assume identical electrodes separations for  $g_1$  and  $g_0$ , respectively for  $C1$  and  $C0$  capacitors. Regarding the total charge density per unit surface  $Q_T$ , it has been assigned values consistent with what is found in MOS and MEMS devices under normal operation (note that  $Q_T$  is a dummy variable used to generate charge densities, related potentials and electric fields in  $C1$  and  $C0$ ). Solving the set of equations involving electrostatics and semiconductors physics gives the charge densities on each capacitor plate, as well as the surface potential in the semiconductor. Next, the ratio of the Coulomb to the energy-based forces  $F_C/F_U$  (see relations (20) and (21)) are displayed in Figures 6a and 6b for the two values of the capacitors gap.

Considering the case where the electrodes gap is 100 nm, Figure 6a confirms that  $F_U$  does not equal  $F_C$ , which was indeed the argument revealing some inconsistency in the “regular” expression of electrostatic energy. This mismatch gets smaller for substrate doping densities greater than  $10^{16} \text{ cm}^{-3}$ . However, it is worth noticing that the link with the extra energy is not straightforward when comparing Figures 5a and 5b with Figures 6a and 6b since a large amount of extra energy does not necessarily mean a large discrepancy in the Coulomb and energy derived forces. In the capacitor system, the dependence of the force with the doping and charge density is less evident as it depends on how this energy is affected by the electrodes separation, i.e. the derivative of the energy with the gap  $g_1$ . In addition, part of the total electric energy is also stored between the plates.

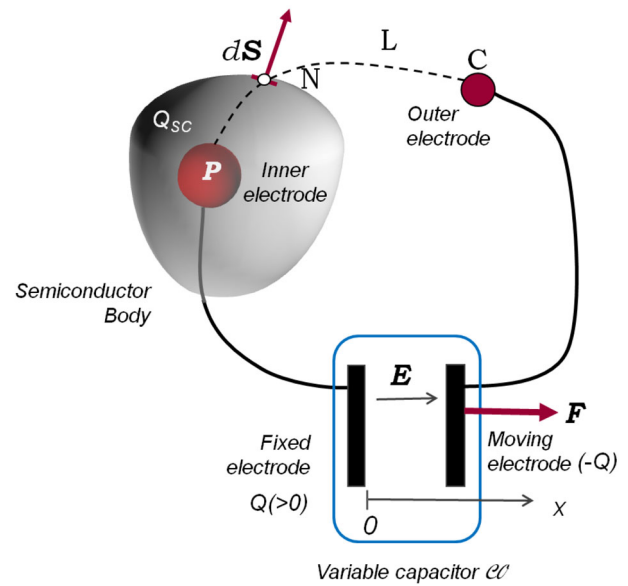
Decreasing the gap down to 10 nm is even more significant in terms of forces deviation. Figure 6b reveals that  $F_C$  and  $F_U$  can differ by nearly one order of magnitude for a doping density of  $10^{15} \text{ cm}^{-3}$ , and by a factor close to 4 for highly doped substrates. The reason why  $F_C/F_U$  is enhanced in smaller gap capacitors can be understood as follows: as far as the energy stored in free space between the plates exceeds the energy stored inside the semiconductor,  $A_{Extra}$  will have a limited influence and  $F_U$  will not depart so much from  $F_C$ . This happens for relatively ‘large’ systems with a gap laying in the micrometer range. But in nanometer scale capacitors, the electric energy stored inside the semiconductor becomes comparable with the electric field energy stored between the plates (in free space). In this case, energy-based and Coulomb forces may deviate significantly from one another. As rule of thumb, we can say that as far as the gap between the plates and the extension of non-neutral regions in the semiconductor are in the same range,  $A_{Extra}$  might not be negligible. This will happen for deep sub-micrometer semiconductor based capacitor systems.



**Fig. 6.** Ratio between the electric force calculated using the Coulomb force ( $F_C$ ) over the force obtained from the standard definition of electric energy ( $F_U$ ), as a function of the charge density, and for different doping concentrations. Two values of electrode separation are addressed, namely 100 nm (a) and 10 nm (b). These calculations reveal that the usual definition of electrostatic energy is not giving the correct value. These discrepancies are even more obvious when decreasing the gap from 100 to 10 nm since the impact of the energy stored in the semiconductor layer becomes enhanced in regard to the energy stored in the free space between electrodes.

## 6 Generalisation to arbitrary geometries

The simple picture of an ideal semi-infinite capacitor system revealed unambiguously the need for an additional electric energy in matter, which has been straightforwardly illustrated for ideal semiconductor layers. Here, we propose to generalize this analysis by considering a three dimensional semiconductor body which is separated from an ideal conductor ‘C’ as shown in Figure 7. Likewise in the former analysis, we connect this system to an ideal capacitor  $C_0$ , but here the idea is to transfer some electric energy from  $C_0$  to the Electrode-Semiconductor system



**Fig. 7.** Representation of the system used to derive the electric energy in a semiconductor body of arbitrary shape. The virtual capacitor is used to generate a continuously varying voltage by means of the work done by an external force on one electrode. The energy transfer between the ideal capacitor and the semiconductor-outer electrode system still reveals that an energy term is missing in the standard formulation of the electrostatic energy.

by increasing gradually the gap between the pre-charged capacitor plates of  $C_0$  until reaching a preset value. Basically, such a process will raise continuously the potential  $V_P$  at node  $P$  in the semiconductor while engendering a transfer of charges between the two systems.

As in the first experiment, this method will involve a mechanical work generated upon moving the plates against the Coulomb force that should still satisfy the principle of energy conservation for the whole system. The advantage for using this setup to generate a potential is that it relies on similar thermodynamics arguments that have been introduced in Section 2 given that both experiments share equal concepts.

### 6.1 Analysis of the energy transfer process

The electric energy stored in the capacitor  $C_0$  can be calculated from relation (1) since we assume ideal conductors for the variable capacitor plates (no penetration of the electric field in conductors):

$$U_{C_0} = \frac{1}{2} \int_{\Omega} \rho \psi d\Omega = \frac{1}{2} Q (V_P - V_C), \quad (56)$$

where  $Q$  is the *total charge* (assumed positive without loss of generality) on the fixed electrode of  $C_0$ , which should not be mistaken with the *charge density* per unit area, and  $V_P$  and  $V_C$  are the inner and outer electrode potentials (see Fig. 7). Since the three dimensional semiconductor

has a limited volume, we must also assign a finite value to  $C_0$  otherwise any *finite* transfer of charges from  $C_0$  to the semiconductor will affect neither the charge density, nor the potential drop across  $C_0$ .

The electric field between the plates and the potential difference are given by (the sign of  $E$  is consistent with the orientation of the  $x$  axis):

$$E = \frac{1}{\epsilon_0} \frac{Q}{S}, \tag{57}$$

$$V_p - V_C = \frac{x}{\epsilon_0} \frac{Q}{S}, \tag{58}$$

where  $S$  is the area of the capacitor plates which are separated by  $x$ . Using (56), the electric field energy stored between the capacitor electrodes is:

$$U_{C_0} = \frac{xQ^2}{2\epsilon_0 S}. \tag{59}$$

As in Section 2, we assume rigid bodies maintained at a fixed temperature and we invoke the same set of hypothesis. The mechanical work spent by applying a force  $\mathbf{F}$  on one electrode to increase the gap between the plates by  $d\mathbf{x}$  is related to the change in the electric energy of the whole system  $dU_E$  (note that the work increases the energy of the system when increasing the gap between the electrodes since  $\mathbf{F} \cdot d\mathbf{x}$  is then positive):

$$\delta W_F = \mathbf{F} \cdot d\mathbf{x} = dU_E = dU_{C_0} + dU_\Omega. \tag{60}$$

Here  $U_\Omega$  holds for the electric energy stored in the semiconductor body and in free space, excluding the capacitor  $C_0$  which is already accounted in  $U_{C_0}$ .

From relation (59), we obtain:

$$\delta W_F = \mathbf{F} d\mathbf{x} = \frac{Q^2}{2\epsilon_0 S} dx + \frac{xQ}{\epsilon_0 S} dQ + dU_\Omega. \tag{61}$$

On the other hand, an attractive Coulomb force exists between the plates. Its module is given by:

$$|F_C| = Q \frac{E}{2} = \frac{Q^2}{2\epsilon_0 S}. \tag{62}$$

Under quasi-static conditions, the applied force  $\mathbf{F}$  and the Coulomb force  $\mathbf{F}_C$  should compensate each other. Therefore the work spent upon moving the plate by  $d\mathbf{x}$  is:

$$\delta W_F = -\mathbf{F}_C d\mathbf{x} = \frac{Q^2}{2\epsilon_0 S} dx. \tag{63}$$

Then, using (61) with (63), we get:

$$dU_\Omega = \frac{-xQ}{\epsilon_0 S} dQ = (V_P - V_C) dQ_{SC}, \tag{64}$$

where  $Q_{SC}$  represents the charge in the semiconductor body and, since the total charge on connected electrodes is invariant (i.e.  $\delta(Q_{SC} + Q) = 0$ ),  $dQ_{SC} = -dQ$ .

The way  $dQ$  and  $dx$  are interrelated depends on the interaction between the semiconductor body and the conductor  $C$ . Here, what matters is that  $V_P - V_C$  can take any

value (provided  $Q$  is not null) by varying the distance between the electrodes. In particular, when electrodes are “in contact” (but still “isolated” electrically), the potential drops to zero. Therefore, using the same kind of topology as in Section 2, we create a virtual variable voltage source while satisfying the principle of energy conservation.

Next, relation (64) can be formally written as a function of the charge transferred to the semiconductor body:

$$dU_\Omega = V_{PC}(Q_{SC}) \cdot dQ_{SC} \tag{65}$$

where  $V_{PC} = V_P - V_C$  depends implicitly on  $Q_{SC}$ .

### 6.2 Derivation of the total electrostatic energy in the semiconductor body

According to Gauss theorem, a variation in the semiconductor charge reverts to a variation in the flux of the surface displacement vector  $d\mathbf{D}$  through the surface  $S_{SC}$ , where  $S_{SC}$  consists of the combination of the outer surface and inner semiconductor-electrode interface (see Fig. 7):

$$dU_\Omega = V_{PC}(Q_{SC}) \left( \oint_{S_{SC}} d\mathbf{D} \cdot d\mathbf{S} \right). \tag{66}$$

Here  $d\mathbf{S}$  is the unitary surface vector oriented outward for the outer surface, and inward for the inner semiconductor-contact interface (note that a virtual “wire” should be created inside the semiconductor body to access the inner contact. However, this can be made infinitesimally narrow so that its shape will not contribute to the surface and volume integrals that come into play with the divergence theorem).

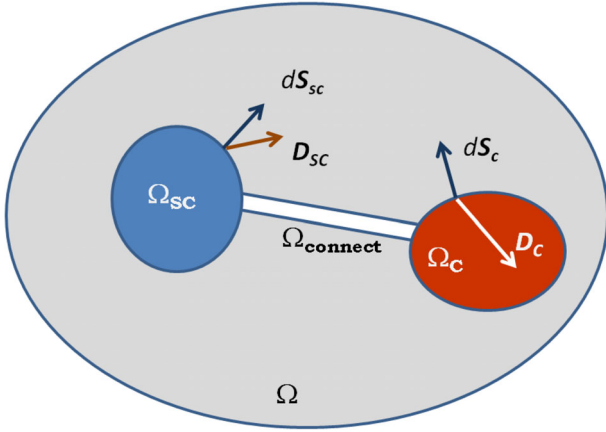
Since the potential  $V_{PC}$  is constant, it can be moved inside the integral:

$$dU_\Omega = \oint_{S_{SC}} V_{PC}(Q_{SC}) (d\mathbf{D} \cdot d\mathbf{S}). \tag{67}$$

Figure 7 shows a line  $L$  joining the point  $P$  in the semiconductor to the external conductor  $C$ . This line crosses the surface  $S$  of the semiconductor at  $N$ . Since the potential drop between  $P$  and  $C$  does not depend on the path (the voltage drop is zero in a closed loop), we can write:

$$dU_\Omega = \oint_{S_{SC}} ((V_P - V_N) + (V_N - V_C)) \cdot (d\mathbf{D} \cdot d\mathbf{S}). \tag{68}$$

In relation (68), the quantity  $(V_N - V_P)$  represents the potential drop in the semiconductor between the inner contact and the point  $N$  at the surface (which reverts to  $\psi_S - \psi_t$  in the case of the semiconductor layer discussed



**Fig. 8.** Topological representation of the three dimensional system involving a semiconductor with an outer electrode. The labeled areas are those used in the divergence theorem.

in Sect. 2):

$$\begin{aligned}
 dU_{\Omega} &= - \oint_{S_{SC}} \psi_S \cdot (d\mathbf{D} \cdot d\mathbf{S}) + \oint_{S_{SC}} (V_N - V_C) \cdot (d\mathbf{D} \cdot d\mathbf{S}) \\
 &= - \oint_{S_{SC}} \psi_S \cdot (d\mathbf{D} \cdot d\mathbf{S}) + \oint_{S_{SC}} V_N \cdot (d\mathbf{D} \cdot d\mathbf{S}) \\
 &\quad - V_C \cdot \oint_{S_{SC}} d\mathbf{D} \cdot d\mathbf{S} \\
 &= - \oint_{S_{SC}} \psi_S \cdot (d\mathbf{D} \cdot d\mathbf{S}) + \oint_{S_{SC}} V_N \cdot (d\mathbf{D} \cdot d\mathbf{S}) \\
 &\quad - V_C \cdot \delta Q_{SC}. \tag{69}
 \end{aligned}$$

The last integral on the right hand side of (69) can be transformed into a field volume integral by using the divergence theorem. To do so, we consider the volume in the whole space (without  $C0$ ), *excluding the semiconductor*, i.e.  $\Omega - \Omega_{SC}$ . This consists in the volume of the semiconductor  $\Omega_{SC}$ , the volume of the electrode  $\Omega_C$  and the volume of the connecting “tube” between them  $\Omega_{Cconnect}$  introduced to ensure the continuity for the inner surface of  $\Omega - \Omega_{SC}$  (see Fig. 8).

In order to simplify relation (69), we start using the divergence theorem:

$$\begin{aligned}
 \int_{\Omega - \Omega_{SC}} \text{div}(V \cdot d\mathbf{D}) d\Omega &= \int_{\Omega - \Omega_{SC}} \text{grad}(V) \cdot d\mathbf{D} \cdot d\Omega \\
 &\quad + \int_{\Omega - \Omega_{SC}} V \cdot \text{div}(d\mathbf{D}) \cdot d\Omega \tag{70}
 \end{aligned}$$

leading to:

$$\begin{aligned}
 \int_{\Omega - \Omega_{SC}} \text{div}(V \cdot d\mathbf{D}) \cdot d\Omega &= - \int_{\Omega - \Omega_{SC}} \mathbf{E} \cdot d\mathbf{D} \cdot d\Omega \\
 &\quad + \int_{\Omega - \Omega_{SC}} V \cdot d\rho \cdot d\Omega. \tag{71}
 \end{aligned}$$

Since there is no charge in  $\Omega - \Omega_{SC}$ , the second integral in (71) vanishes, leading to:

$$\int_{\Omega - \Omega_{SC}} \text{div}(V \cdot d\mathbf{D}) \cdot d\Omega = - \int_{\Omega - \Omega_{SC}} \mathbf{E} \cdot d\mathbf{D} \cdot d\Omega. \tag{72}$$

Noting that  $\Omega - \Omega_{SC}$  on one side, and  $\{\Omega_{SC}, \Omega_C\}$  on the other side have opposite surface orientations, using the divergence theorem again and noting that at infinity the displacement vector must vanish, we have:

$$\begin{aligned}
 \int_{\Omega - \Omega_{SC}} \text{div}(V \cdot d\mathbf{D}) \cdot d\Omega &= - \oint_{S_{SC}} V_N \cdot (d\mathbf{D} \cdot d\mathbf{S}) \\
 &\quad - \oint_{S_C} V_C \cdot (d\mathbf{D} \cdot d\mathbf{S}) \\
 &\quad - \oint_{S_{Connect}} V \cdot (d\mathbf{D} \cdot d\mathbf{S}). \tag{73}
 \end{aligned}$$

Next, the connecting element is made infinitesimally small, so that the last integral over  $S_{Connect}$  in (73) can be omitted:

$$\begin{aligned}
 \int_{\Omega - \Omega_{SC}} \text{div}(V \cdot d\mathbf{D}) \cdot d\Omega &= - \oint_{S_{SC}} V_N \cdot (d\mathbf{D} \cdot d\mathbf{S}) \\
 &\quad - \oint_{S_C} V_C \cdot (d\mathbf{D} \cdot d\mathbf{S}). \tag{74}
 \end{aligned}$$

Noting that  $V_C$  is constant on the conductor surface and that the integral of  $d\mathbf{D}$  over the conductor surface  $S_C$  is simply  $-dQ_{SC}$ , as imposed by the charge neutrality principle, we can write:

$$\int_{\Omega - \Omega_{SC}} \text{div}(V \cdot d\mathbf{D}) \cdot d\Omega = - \oint_{S_{SC}} V_N \cdot (d\mathbf{D} \cdot d\mathbf{S}) + V_C \cdot \delta Q_{SC}. \tag{75}$$

Using the above identity in relation (69), the incremental energy transfer from the capacitor to the rest of the system, including free space (except  $C0$ ) is:

$$dU_{\Omega} = - \oint_{S_{SC}} \psi_S (d\mathbf{D} \cdot d\mathbf{S}) - \int_{\Omega - \Omega_{SC}} \text{div}(V \cdot d\mathbf{D}) \cdot d\Omega. \tag{76}$$

Next, the total work exchanged is obtained by integrating (76) over the displacement vector while noting that there is no charge in  $\Omega - \Omega_{SC}$ :

$$U_{\Omega} = - \int_{\mathbf{D}} \oint_{S_{SC}} \psi_S \cdot d\mathbf{S} \cdot d\mathbf{D} + \int_{\mathbf{D}} \int_{\Omega - \Omega_{SC}} \mathbf{E} \cdot d\Omega \cdot d\mathbf{D}. \tag{77}$$

Permuting the integration sequence, we obtain:

$$U_{\Omega} = - \oint_{S_{SC}} \int_{\mathbf{D}} \psi_S \cdot d\mathbf{D} \cdot d\mathbf{S} + \int_{\Omega - \Omega_{SC}} \int_{\mathbf{D}} \mathbf{E} \cdot d\mathbf{D} \cdot d\Omega. \quad (78)$$

The electric energy supplied by the virtual voltage source is divided in some energy stored in free space and in the semiconductor body. On the other hand, the electrostatic energy in the free space  $\Omega$  (excluding  $C0$ , but including the semiconductor) is still given by relation (3).

Likewise for the analysis of the capacitor plate, the difference in these quantities represents the additional free energy for the body:

$$\begin{aligned} A_{Extra} &= U_{\Omega} - U_E \\ &= - \oint_{S_{SC}} \int_{\mathbf{D}} \psi_S \cdot (d\mathbf{D} \cdot d\mathbf{S}) - \int_{\Omega_{SC}} \int_{\mathbf{D}} \mathbf{E} \cdot d\mathbf{D} \cdot d\Omega. \end{aligned} \quad (79)$$

Therefore, the new definition of the electric energy of a system including an arbitrary semiconductor body is proposed:

$$\begin{aligned} U_E^{new} &= \int_{\Omega} \int_{\mathbf{D}} \mathbf{E} \cdot d\mathbf{D} \cdot d\Omega + A_{Extra} \\ &= \int_{\Omega - \Omega_{SC}} \int_{\mathbf{D}} \mathbf{E} \cdot d\mathbf{D} \cdot d\Omega - \oint_{S_{SC}} \int_{\mathbf{D}} \psi_S \cdot (d\mathbf{D} \cdot d\mathbf{S}). \end{aligned} \quad (80)$$

It is straightforward to see that relation (80) is the generalization of relation (32) that was developed for a planar geometry. Therefore, even though we followed a quite different approach, we conclude that an additional free energy created by an electric field is predicted in semiconductors, and very likely in matter in general, and only in some special cases this energy is not engendered.

We can proceed further these developments and propose an expression in terms of the local charge density. Using the divergence theorem for the surface integral of (80), we obtain:

$$\begin{aligned} \oint_{S_{SC}} \int_{\mathbf{D}} \psi_S \cdot (d\mathbf{D} \cdot d\mathbf{S}) &= \int_{\Omega_{SC}} \int_{\mathbf{D}} \text{div}(\psi \cdot d\mathbf{D}) \cdot d\Omega \\ &= - \int_{\Omega_{SC}} \int_{\mathbf{D}} (\mathbf{E} \cdot d\mathbf{D}) \cdot d\Omega \\ &\quad + \int_{\Omega_{SC}} \int_{\rho} (\psi \cdot d\rho) \cdot d\Omega. \end{aligned} \quad (81)$$

Introducing this identity in relation (79), we find that the energy contribution coming from the semiconductor body takes a very simple form:

$$A_{Extra} = - \int_{\Omega_{SC}} \int_{\rho} (\psi d\rho) \cdot d\Omega. \quad (82)$$

At least in semiconductors,  $A_{Extra}$  is always positive, i.e. the internal energy is increased upon generation of non-neutral regions.

In addition, using relation (1) (valid only if the system is linear [5]) for the standard expression of electric energy, we can write the new energy as a function of the potential and charge density only:

$$U_E^{new} = \frac{1}{2} \int_{\Omega_{\infty}} (\phi\rho) \cdot d\Omega - \int_{\Omega_{SC}} \int_{\rho} (\psi d\rho) \cdot d\Omega. \quad (83)$$

Two clarifications need to be done. Firstly, the volume  $\Omega$  in (80) is used to identify the system of interest (semiconductor body with outer electrode) excluding the ‘capacitor voltage source’. But this still represents the whole space termed as  $\Omega_{\infty}$  in relation (1), reason why it is used in place of  $\Omega$  in relation (83). Next, the meaning of the potentials  $\phi$  and  $\psi$  should be well understood. Whereas  $\phi(\mathbf{x})$  is the potential created at  $\mathbf{x}$  by all the charges in the sense given by relation (1) and should therefore vanish at infinity,  $\psi(\mathbf{x})$  is the difference in the potentials between the inner electrode  $\mathbf{P}$  and a coordinate  $\mathbf{x}$  inside the semiconductor body (Fig. 7).

Finally, when considering isotropic media and introducing the polarization vector as in reference [2], the different contributions to the electrostatic energy are highlighted:

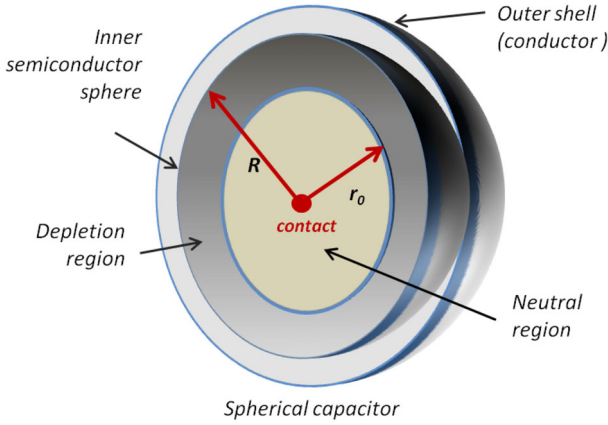
$$\begin{aligned} U_E^{new} &= \varepsilon_0 \cdot \int_{\Omega} \int_{\mathbf{E}} (\mathbf{E} \cdot d\mathbf{E}) \cdot d\Omega + \int_{\Omega_{SC}} \int_{\mathbf{P}} (\mathbf{E} \cdot d\mathbf{P}) \cdot d\Omega \\ &\quad - \int_{\Omega_{SC}} \int_{\rho} (\psi \cdot d\rho) \cdot d\Omega. \end{aligned} \quad (84)$$

The first term represents the total energy stored in the electric field in the whole space, including the semiconductor body. The second term represents the energy supplied to polarize matter [1,2]. Lastly, the third term represents the new energy  $A_{Extra}$ .

Different expressions for the electrostatic energy with the new definition are then obtained. Note that whereas relations (80) and (84) generalize relation (3), relation (83) is somewhat more restrictive as it makes use of relation (1) and hence should be used with the same restriction, i.e. for linear polarizable systems (in the sense discussed in Ref. [5]).

### 6.3 The full depletion approximation in a semiconductor sphere as a case of study

We illustrate the use of generalized expressions of the extra energy in the case of a p-type doped semiconductor sphere of radius  $R$  where the full depletion approximation is invoked. We consider a kind of semiconductor-capacitor system as shown in Figure 9. Likewise for the planar geometry discussed formerly, applying a lower potential to the silicon with respect to the surrounding conductive sphere will deplete the surface of the



**Fig. 9.** Cross section of the spherical semiconductor body surrounded by an ideal counter electrode that creates a depletion in the inner sphere (electrical potentials are applied at the center and on the outer conductive shell). The radius of the semiconductor sphere is  $R$  and the radius where the transition between *full depletion-neutral semiconductor* takes place is labeled  $r_0$ .

semiconductor sphere from mobile carriers. In addition, as the semiconductor is uniformly doped, there is no angular dependence of any physical quantity, implying that the Poisson equation in spherical coordinate satisfies:

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) \psi(r) = \frac{qN_A}{\epsilon_{SC}}. \quad (85)$$

The full depletion approximation supposes that the core of the semiconductor sphere remains neutral up to a radius  $r_0$ , then is fully depleted from  $r_0$  to  $R$  ( $\psi(r_0) = 0$ ,  $E(r_0) = 0$ ). Imposing these conditions to relation (85), the potential and the radial electric field depend on the radius  $r$  and on the parameter  $r_0$  such as:

$$\psi(r) = \frac{qN_A}{6\epsilon_{SC}} r^2 + \frac{qN_A}{3\epsilon_{SC}} r_0^3 \frac{1}{r} - \frac{qN_A}{2\epsilon_{SC}} r_0^2, \quad (86)$$

$$E(r) = \frac{qN_A}{3\epsilon_{SC}} \left( \frac{r_0^3}{r^2} - r \right). \quad (87)$$

We are now able to evaluate the extra energy defined from relation (79) for instance.

Concerning the surface integral, since we are considering a sphere with no angular dependence, the integral over the surface for a given surface displacement vector  $D_S$  simplifies as:

$$\oint_S \int_D \psi_S \cdot (d\mathbf{D} \cdot d\mathbf{S}) = 4\pi R^2 \int_D \psi_S \cdot d\mathbf{D}. \quad (88)$$

Concerning the surface potential, this is straightforwardly obtained from relation (86):

$$\psi_S = \psi(R) = \frac{qN_A}{6\epsilon_{SC}} R^2 + \frac{qN_A}{3\epsilon_{SC}} r_0^3 \frac{1}{R} - \frac{qN_A}{2\epsilon_{SC}} r_0^2. \quad (89)$$

Next, the incremental surface displacement vector is recovered from relation (87):

$$dD(R) = qN_A \left( \frac{r_0^2}{R^2} \right) dr_0 \quad (90)$$

(note that there is no dependence on the radius since  $D_S$  is defined at the surface only).

After calculations, the surface integral becomes:

$$\begin{aligned} \oint_S \int_D \psi_S (d\mathbf{D} d\mathbf{S}) &= \frac{-4\pi}{\epsilon_{SC}} (qN_A)^2 \\ &\times \int_{r_0}^R \left( \frac{R^2 r_0^2}{6} + \frac{1}{3} \frac{r_0^5}{R} - \frac{r_0^4}{2} \right) dr_0 \\ &= \frac{-2\pi}{9\epsilon_{SC}} (qN_A)^2 \\ &\times \left[ \frac{R^5}{5} - R^2 r_0^3 - \frac{r_0^6}{R} + \frac{9}{5} r_0^5 \right]. \quad (91) \end{aligned}$$

Next, the second contribution is evaluated.

Still assuming that there is no angular dependence and that the medium is linear, we have ( $d\Omega = 4 \times \pi r^2 dr$ ):

$$\begin{aligned} \int_{\Omega_{SC}} \int_D \mathbf{E} \cdot d\mathbf{D} \cdot d\Omega &= \frac{1}{2} \cdot \epsilon_{SC} \cdot \int_{\Omega_{SC}} \mathbf{E}^2 \cdot d\Omega \\ &= \frac{2 \cdot \pi}{9 \cdot \epsilon_{SC}} \cdot (q \cdot N_A)^2 \cdot \int_{r_0}^R \left( \frac{r_0^3}{r} - r^2 \right)^2 \cdot dr \\ &= \frac{2 \cdot \pi}{9 \cdot \epsilon_{SC}} \cdot (q \cdot N_A)^2 \cdot \left( \frac{R^5}{5} - R^2 \cdot r_0^3 - \frac{r_0^6}{R} + \frac{9}{5} \cdot r_0^5 \right). \quad (92) \end{aligned}$$

It is worth noticing that in (91), integration is performed over the parameter  $r_0$ , which is the radius where depletion starts, whereas in (92) the integration is done over the variable  $r$  from  $r_0$  to  $R$ . Comparing both terms reveals that these integrals are equal and will cancel each others in relation (79). Therefore, the regular definition of the electrostatic energy remains valid when the full depletion approximation is used even in case of the semiconductor spheres.

So far we have carried out the analysis making of use of fields, which was instructive to understand how these relationships can be used in practical cases. In the next section, we will adopt the dual charge-potential approach.

#### 6.4 The full depletion approximation in arbitrary geometries

As stated in relation (82), an alternative expression of  $A_{Extra}$  in terms of charges and potentials is possible. Here, we can gain more insight when integrating by parts the

inner integral of relation (82). This gives:

$$A_{Extra} = - \int_{\Omega_{SC}} \left( \psi(\mathbf{r}) \cdot \rho(\mathbf{r}) \Big|_i^f - \int_{\psi_i}^{\psi_f} \rho(\mathbf{r}) \cdot d\psi(\mathbf{r}) \right) d\Omega, \quad (93)$$

where the subscript ‘i’ and ‘f’ hold for initial and final values of the physical quantities evaluated at the same coordinate  $\mathbf{r}$ .

Now, if the charge density  $\rho(\mathbf{r})$  at  $\mathbf{r}$  does not depend on the potential  $\psi(\mathbf{r})$ , as it may happen insulators or in semiconductors when the full depletion approximation holds, we have:

$$\begin{aligned} A_{Extra} &= - \int_{\Omega_{SC}} \left( \psi(\mathbf{r}) \Big|_i^f \cdot \rho(\mathbf{r}) - \rho(\mathbf{r}) \cdot \int_{\psi_i}^{\psi_f} d\psi(\mathbf{r}) \right) d\Omega \\ &= - \int_{\Omega_{SC}} \left[ \psi(\mathbf{r}) \Big|_i^f \cdot \rho(\mathbf{r}) - \rho(\mathbf{r}) \cdot \psi(\mathbf{r}) \Big|_i^f \right] d\Omega. \end{aligned} \quad (94)$$

It is clear that the integrand is null at each coordinate where this assumption holds. Then, if this condition is verified in the whole body, the integration over  $\Omega_{SC}$  is also null.

This result generalizes the preliminary conclusions carried out in former sections for planar and spherical geometries. It confirms that whatever the shape, no extra energy is needed when dealing with neutral or charged insulators, as well as with semiconductors when adopting the full depletion approximation assumption.

## 7 Conclusion

Following ground theoretical developments, we conclude on the necessity to introduce a new electric energy contribution in semiconductors, and possibly in a

variety of materials, which, to the best of the author’s knowledge, has never been evidenced before and is still absent from text books. Indeed, depending on the magnitude and orientation of the external electric field, such energy may largely exceed the well-known free energy related to dipoles polarization. A generalization to three-dimensional systems is proposed and general rules regarding the need for such a correction are discussed. It comes out that under special situations, as for instance in insulators or when the full depletion approximation holds, this energy vanishes and gives back the commonly accepted definition of electric energy. Besides this fundamental aspect, we anticipate that totally neglecting this new term is not consistent with the energy conservation principle. Therefore, it seems mandatory to rework some definition of electric energy in matter in order to reconcile electric energy with basics of electrostatics.

## References

1. S. Bobbio, *Electrodynamics of materials; forces, stresses and energies in solids and fluids* (Academic Press, New York, 1999)
2. S. Bobbio, Eur. Phys. J. B **16**, 43 (2000)
3. L.D. Landau, E.M. Lifshitz, *Electrodynamique des Milieux Continus* (MIR, Moscou, 1990)
4. Y. Zimmels, J. Magn. Mater. **292**, 433 (2005)
5. J.D. Jackson, *Classical Electrodynamics*, 3rd edn. (John-Wiley and Sons, 1999)
6. Y. Zimmels, Phys. Rev. E **65**, 036146 (2002)
7. D.R. Frankl, Surf. Sci. **9**, 73 (1968)
8. S.M. Sze, *Physics of semiconductor devices*, 2nd edn. (John Wiley & Sons, 1981)

**Open Access** This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.