Three- to two-dimensional transition in electrostatic screening of point charges at semiconductor surfaces studied by scanning tunneling microscopy

A. Laubsch,[*] K. Urban, and Ph. Ebert[†]

Institut für Festkörperforschung, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
(Received 13 September 2009; published 14 December 2009)

The electrostatic screening of localized electric charges on semiconductor surfaces is investigated quantitatively by statistically analyzing the spatial distribution of thermally formed positively charged anion surface vacancies on GaAs and InP(110) surfaces. Two screening regimes are found: at low vacancy concentrations the vacancy charges are found to be three-dimensionally screened by bulk charge carriers. The corresponding screening length, which increases strongly with decreasing carrier concentration, is best described by the classical bulk screening length evaluated with a surface dielectric constant. With increasing vacancy concentration at given bulk carrier concentration, a three- to two-dimensional screening transition occurs. At high vacancy concentrations, the screening is found to be governed by charge carriers located in a two-dimensional surface vacancy defect band, which is partially filled due to the vacancy-induced surface band bending.

I. INTRODUCTION

The ability to reproducibly incorporate charged dopants and defects with accurately controlled concentrations and spatial distributions is of utmost importance for technological applications of semiconductor materials. In the light of the ongoing miniaturization of semiconductor devices, this task is even more critical, since the misplacement of an individual charged dopant or defect can ruin future nanoscale semiconductor devices. However, the nanoscale distribution of impurities and defects is still mostly beyond technical control. At nanometer scales, the intrinsic interactions between defects and/or dopant atoms ultimately govern the incorporation and the spatial distribution within the device. Particularly important in this respect is the interaction of *charged* defects at semiconductor interfaces and surfaces.

The generally accepted view is that the charge of a defect is screened by a spatial redistribution of free charge carriers, resulting in screened Coulomb potential interactions between the charge centers.[1] Although this fundamental principle of screening by free charge carriers has been derived theoretically since long for the bulk[1–3] and low-dimensional systems,[4,5] and more recently for surfaces,[6–8] it is experimentally rather difficult to probe reliably the screened Coulomb potential with atomic resolution.

Thus far the screening at surfaces could only be probed by scanning tunneling microscopy (STM),[9–11] due to its atomic resolution combined with an indirect potential sensitivity. However, the experimental results are limited in many respects. First, the screening potentials are frequently approximated by height profiles taken in constant-current STM images.[9–14] Second, the close proximity of the STM tip near the surface induces localized band bendings,[15–18] which modify the potential around charged defects significantly and thus the screening cannot be probed quantitatively. Third, the limited signal to noise ratio of barrier height or surface photovoltage measurements[19–21] restrict a quantitative data analysis. These problems drastically limit the availability of reliable experimental data on screening at semiconductor surfaces. Furthermore, the few existing data are even in open

conflict with theoretical expectations.[14] The currently published experimental values for the screening length (which determines the extend of the screened Coulomb potential) at GaAs and InP(110) surfaces are constantly between 1 and 2 nm independent of the defect and doping concentrations investigated.[9–14] This disagrees with theoretical expectations, which predict a strong increase in the screening length with decreasing carrier concentration. Despite theoretical efforts to clarify this inconsistency by assuming additional screening by defect-induced carriers at the surface,[7,8] no understanding of the apparently short screening length at negligibly low defect concentrations was achieved. In view of the importance of screening in nanoscale devices and as fundamental physical principle, it is, however, crucial to obtain reliable atomically resolved experimental data of the screening potentials, in order to unravel, which physical model(s) describe(s) best the screening at semiconductor surfaces.

Therefore, we probed experimentally the screened Coulomb potential on InP and GaAs cleavage surfaces in a wide range of doping and defect concentrations using a statistical approach, which avoids any tip influence on the data.[14] This allows us to identify a three-dimensional (3D)- to two-dimensional (2D) transition of the screening with increasing vacancy concentration, thereby resolving the previous discrepancies. Our results enable us to explain the different screening effects at surfaces and to provide a comprehensive theoretical explanation.

II. EXPERIMENTAL DETAILS

Screening is an electrostatic process that happens to any charged impurity or defect in the same way. The problem can thus be reduced to understand the screening of a charged center of arbitrary nature. Here we use as model system charged surface anion vacancies, since their concentration can be adjusted reproducibly over a wide range by suitable thermal treatment of *p*-doped GaAs and InP(110) surfaces.[22] The vacancies were investigated by STM with atomic resolution. Since STM measures the electron density of the sample surface and thereby is sensitive to a local screened

Coulomb potential,[23] one could, in principle, try to directly image the screening-charge cloud surrounding each charged vacancy[11–13,24] and quantify it following the method presented in Ref. 25 in absence of a tip-induced band bending. However, for the carrier concentrations present in the samples investigated here, the metal tip of the microscope induces a spatially confined band bending at the surface,[15–18] which perturbs the screening potential we want to measure.

Therefore, we rather use tools from statistical mechanics that are independent of individual imaging conditions and the close proximity of the biased tip of the STM. Analyzing *only the vacancies' spatial distribution* on the sample surface, the screened potential around a charged vacancy can be deduced by evaluating the pair-correlation function $c(\vec{r})$ of the vacancies.[14]

First, we deduced from the STM images the positions of all vacancies and calculated the distances between all possible vacancy pairs. This gives us the measured probability distribution of pair distances.

Second, we obtain the 2D pair-correlation function $c(\vec{r})$ by normalizing the measured probability distribution of pair distances to the numbers of all geometrically possible pair configurations for each separation and to the vacancy concentration. $\vec{r}$ is the 2D distance vector along the surface, i.e., in our case a surface lattice vector.

Third, the pair-correlation function $c(\vec{r})$ is related to the so-called mean force potential $W(\vec{r})$, through $W(\vec{r}) = -kT \times \ln[c(\vec{r})]$.[26] In the low-density limit the mean force potential $W(\vec{r})$ equals the interaction energy $V(\vec{r})$, which is given here by the component of the screened Coulomb potential along the surface. The strength and extent of the interaction potential affects this low-density limit. In the present case the low-density limit it is valid for all investigated vacancy concentrations, because the screening lengths are significantly shorter than the average vacancy-vacancy separation.[27] If we would use, for example, the more sophisticated hypernetted chain closure,[28] we obtain within 1 meV the same pair-interaction potential. This is well below the experimental accuracy and thus we do not need to take interactions between more than two vacancies into account (i.e., many-body interactions[29]). Therefore, we use the so-called Boltzmann-closure equation yielding for the pair-interaction energy

$$V(\vec{r}) = -kT \times \ln[c(\vec{r})]. \tag{1}$$

$V(\vec{r})$ is given by the screening potential, since we investigate the interaction between the charged vacancies. Only at very short distances smaller than 2 lattice constants, atomic lattice relaxation effects modify the repulsive interaction.[14] Here we do not consider these very small separations, since for such separations the screening approach used is anyway not valid.

The statistical approach requires that the vacancy distribution at the surface is sufficiently close to thermal equilibrium. We obtained such vacancy distributions using in ultrahigh vacuum cleaved surfaces of different *p*-type GaAs and InP crystals (pressure $< 1 \times 10^{-8}$ Pa). The bulk doping concentrations of the crystals ranged between about 1×10^{18} and 3×10^{20} cm^{-3}. After cleavage the samples were subjected to different annealing temperatures up 480 K and 2500 min in

ultrahigh vacuum to obtain a set of different high vacancy concentrations in thermal equilibrium. After radiation cooling the samples were directly transferred to the STM and investigated. Other samples were left at room temperature in ultrahigh vacuum, where vacancies also form thermally, however, with lower rates.[22] In this way small vacancy concentrations are accessible to probe the low-density interaction. Note, that at all temperatures, vacancies are formed thermally and thus the vacancies have also sufficient thermal energy to overcome the diffusion barriers and reach an equilibrium distribution for the given vacancy concentration and vacancy interaction. During annealing only anion vacancies form on the surfaces. The concentrations of all other defects are negligible and thus do not affect the obtained data. Furthermore, we use imaging conditions as discussed in Ref. 30, where the tip does not induce vacancy migration events[12,31–33] or produce/modify defects.[24,34]

III. RESULTS

A. Experimental determination of the interaction potential

Figure 1 shows STM images of the vacancy distributions on two cleavage surfaces prepared as outlined above. Along the $[1\bar{1}0]$ direction the atomic rows consisting of alternating anion (As/P) and cation (Ga/In) atoms are visible. Since the images were both measured at negative voltages, only the occupied dangling bonds above the anions contribute to the atomic contrast.[35] Therefore, both STM images can be interpreted as images of the anion sublattice. Two types of local contrast are visible in the STM images: bright contrasts (marked dopant) arise from negatively charged Zn dopant atoms incorporated on cation lattice sites.[12,36,37] The negative charge localized at the dopant atom is screened by a redistribution of the free holes. This leads to a hole accumulation and thereby to a local upward band bending due to the presence of the resulting screened Coulomb potential. The local band bending leads to a locally increased tunnel current, which is compensated by an increased tip-sample separation in the constant-current mode used for measuring the STM images. This increased tip-sample separation is imaged as brighter contrast. In analogy the local dark contrasts arise from positive charges localized at surface As [Fig. 1(a)] and P vacancies [Fig. 1(b)].[38,39] Sometimes a vacancy appears less dark. This effect is due to nearby dopant atoms, which compensate the charge of the vacancy as shown in detail in Ref. 36. Dopant atoms are only visible at low vacancy concentrations [Fig. 1(a)], since with increasing vacancy concentration, the probability of close vacancy-dopant pairs increases and leads to the compensation of the dopant charge[36] as in case of Fig. 1(b).

In the next step we identified the atomic positions of every vacancy visible in the STM images. From these 2D distributions of vacancy positions, we calculated the pair-correlation function as outlined in the preceding section.

B. The interaction potential for high vacancy concentrations

1. Experimental results

First we focus on the high vacancy concentration data, such as the example shown in Fig. 1(b). Figure 2(a) shows
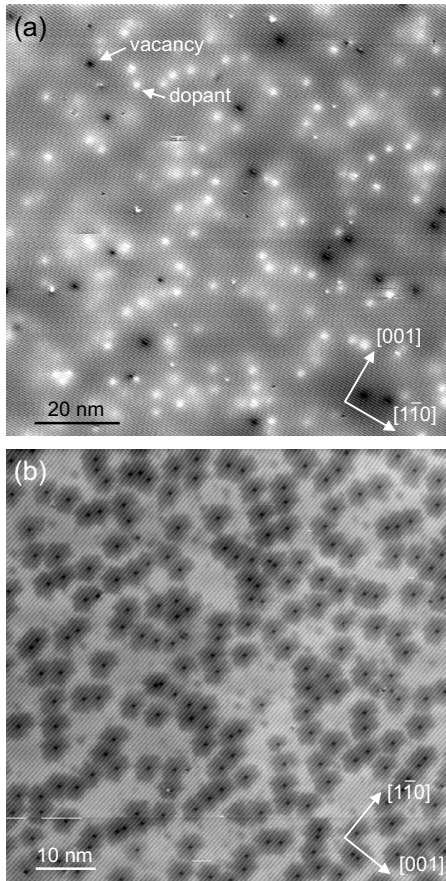
FIG. 1. Constant-current scanning tunneling microscopy images measured at room temperature of (a) a p-doped GaAs(110) cleavage surface after 11 790 min at 286 K and (b) a p-doped InP(110) cleavage surface after 2500 min annealing at 480 K. The average vacancy concentrations for the samples from which these images originate were (a) 0.058% and (b) 1.0%. The images were measured at voltages of (a) -2.5 V and (b) -2.6 V.

the 2D pair-correlation function $c(\vec{r})$ obtained for a P vacancy concentration of 1.0% reached by annealing a p-doped InP(110) surface for 2500 min at 480 K. $\vec{r}$ is a *surface* lattice vector originating at a vacancy site. The pair correlation has values smaller than one (gray colored area) up to 4 nm away from the vacancy site. At larger separations the values of the pair-correlation function are close to 1. This means that close to a vacancy, the probability of finding another vacancy is strongly reduced, whereas at large separations no correlation exists. Furthermore, Fig. 2(a) illustrates that the vacancy-vacancy interaction is circular and thus independent of the atomic symmetry of the surface. Therefore, one can treat the pair-correlation function as a function of the radial distance from the vacancy site defined as $r=|\vec{r}|$ as shown in Fig. 2(b). Again the reduced probability at small separations is visible, whereas at large distances no correlation exists.

The physical origin of this shape of pair-correlation function lies in the repulsive Coulomb interaction between the charged vacancies. In order to extract the pair-interaction energy of the charged vacancies, which corresponds to the screened Coulomb potential surrounding each vacancy, we
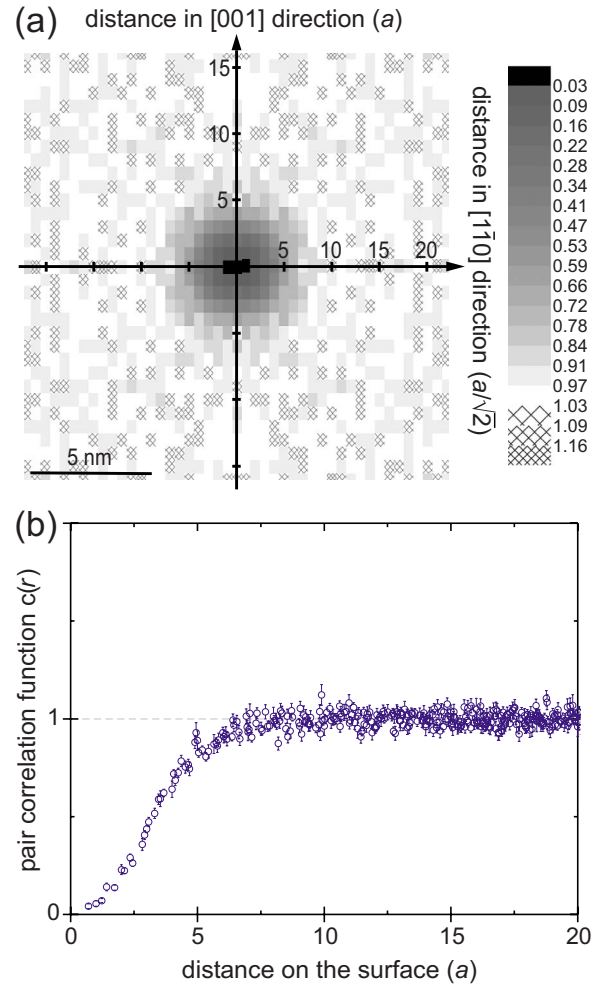

FIG. 2. (Color online) (a) Two-dimensional pair-correlation function $c(\vec{r})$ for positively charged P vacancies on p-doped InP(110) surfaces. The data set shown was obtained from vacancy distributions obtained after annealing at 480 K for 2500 min. The directions of the axes correspond to the high symmetry surface directions. The lattice constant is $a=0.587$ nm. $\vec{r}$ is a surface lattice vector originating at a vacancy site. (b) Pair-correlation function shown as a function of the radial distance $r=|\vec{r}|$. For this data set 48 916 vacancy positions digitized with atomic precision were used. The vacancy concentration of this data set was 1.0%.

used the Boltzmann closure in Eq. (1). This yielded the pair-interaction energy shown in Fig. 3 as open circles.

2. Screening potentials

In order to analyze the pair-interaction energy we turn to the theory of electrostatic screening of localized charges by free charge carriers. In the bulk and at the surface, localized charges are screened by a redistribution of the free charge carriers, which thereby form a screening cloud. In the *bulk* this modifies the Coulomb potential by adding an exponential screening term yielding a screened Coulomb potential with the functional form of a Yukawa potential, i.e., $V_b(R) = q/(4\pi\varepsilon_0\varepsilon_r R)\exp(-R/R_b)$.[1] The screening is in first approximation isotropic depending only of the radial separation R in the bulk. q is the localized charge and R_b is the so-called
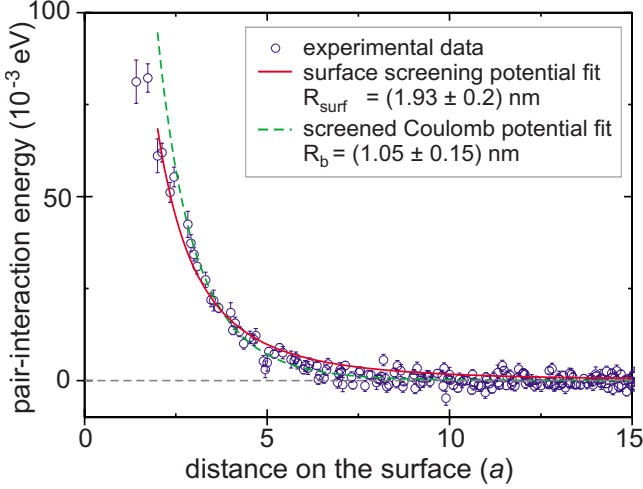
FIG. 3. (Color online) Pair-interaction energy $V(r)$ extracted from the pair-correlation function $c(\vec{r})$ shown in Fig. 2 using the Boltzmann-closure Eq. (1). The experimental data is shown as open circles. The solid line shows a fit to the data using a two-dimensional surface screening model [Eq. (4)]. The dashed line is a fit using a three-dimensional bulklike screened Coulomb potential [Eq. (2)].

bulk screening length, defined by the material's properties, such as dielectric constant ε_r, carrier concentration, and effective mass m_{eff} (see below).

At a surface the spherical symmetry is interrupted and two types of screening were theoretically derived: (i) a bulk (3D) screening of a surface defect[6] and (ii) a purely surface (2D) screening of a surface defect.[4,7] In the following we only consider the screening potential in the surface plane and do not discuss its extension into subsurface layers, because the vacancies are only located at the surface and thus interact only according to the surface component of the interaction potential.

(i) If a surface defect with charge q is screened by free charge carriers in the isotropic bulk and no additional screening charges are present at the surface, the screening potential along the surface plane has the functional form of a screened Coulomb potential[6]

$$V_{b,surf}(r) = \frac{q}{4\pi\varepsilon_0\varepsilon_{surf}}\frac{1}{r}e^{-r/R_b} \qquad (2)$$

with surface dielectric constant $\varepsilon_{surf} = (\varepsilon_r + 1)/2$ being the average of the two media separated by the surface (i.e., the semiconductor and vacuum).[40] The screening length in the surface plane is given by the bulk screening length R_b (Ref. 1)

$$R_b = \sqrt{\frac{2\pi^2\varepsilon_0\varepsilon_r\hbar^3}{e^2 m_{eff}^{3/2}(2\pi kT)^{1/2}F_{-1/2}\left(\frac{E_F}{kT}\right)}}. \qquad (3)$$

$F_{-1/2}$ is the Fermi-Dirac integral as defined in Ref. 1. E_F is the Fermi energy defined by the charge-carrier concentration in the semiconductor. k is the Boltzmann constant and T the temperature of the system. e is the electron charge. The main

feature of Eq. (3) is that the screening length at the surface increases strongly with decreasing carrier concentration (due to the E_F dependence) and can thereby reach very large values.

(ii) In case of only surface screening charges (and no bulk screening) the screening potential is given by[41]

$$V_{surf}(r) = \frac{q}{4\pi\varepsilon_0\varepsilon_{surf}}\frac{1}{r}\left[1 - \frac{\pi}{2}\frac{r}{R_{surf}}\left(H_0\left(\frac{r}{R_{surf}}\right) - N_0\left(\frac{r}{R_{surf}}\right)\right)\right]. \qquad (4)$$

H_0 and N_0 are the zeroth-order Struve and Neumann functions, respectively. r is again the distance from the charge along the surface.

The surface screening length R_{surf} was derived in analogy to the Ref. 1. We obtained

$$R_{surf} = \frac{2\pi\varepsilon_0\varepsilon_{surf}\hbar^2}{g_v e^2 m_{eff}F_{-1}\left(\frac{E_F - E_0}{kT}\right)}$$

$$= \frac{2\pi\varepsilon_0\varepsilon_{surf}\hbar^2}{g_v e^2 m_{eff}}(1 + e^{-(E_F - E_0)/kT}). \qquad (5)$$

E_0 corresponds to the energy level, which governs the occupation of the two-dimensional surface state, where the screening charge carriers are located. g_v is the orbital degeneracy of the band. In our case it can be set to 1, since as shown below the 2D surface band is not related to the materials band structure, but rather a defect band. $F_{-1}(\eta)$ can be defined by $d/d\eta[F_0(\eta)] = F_{-1}(\eta)$.[42]

3. Analysis of the experimental interaction energy

The two screening scenarios discussed above are the limits of the general case, where both surface and bulk-related free charge carriers contribute to the screening. We first compare the two limits with the experimentally measured interaction potential by fitting Eqs. (2) and (4) to the experimental data [dashed (green) and solid (red) lines in Fig. 3].[43] As fit parameters we used the respective screening lengths only. We assumed a charge of $q = +1e$ for the P [As] vacancies on p-doped InP(110) [GaAs(110)] surfaces in accordance with the experimental charge determination[36] and the theoretically calculated charge.[44,45]

At first view both curves provide an equally good fit to the data. Thus a distinction between both screening mechanism is not possible purely on the fit quality. However, if we turn to the obtained screening lengths, clear differences appear. First the screened Coulomb potential fit yields a screening length of (1.05 ± 0.15) nm. According to Eqs. (2) and (3) this value should correspond to the bulk screening length. For a carrier concentration of $(1.3-2.1)\times 10^{18}$ cm^{-3} in the InP sample, the bulk screening length should be 3–4 nm. If we also consider that the vacancies at the surface compensate the dopants and thereby reduce the carrier concentration in the vicinity of the surface by several orders of magnitude (as discussed below), we should expect a bulk screening length larger than 100 nm, i.e., in any case much larger than that

obtained through the screened Coulomb potential fit.

The situation is different for the surface screening potential model. The screening length of (1.93 ± 0.2) nm obtained from the fit is in agreement with the expected screening length [Eq. (5)] in the range of only a few nanometers. If we assume screening charges at the surface in the order of the vacancy concentrations and effective masses $m_{eff} \approx m_0$, with m_0 being the electron mass, we obtain surface screening lengths in the range of 0.5–2 nm. This is in good agreement with the experimental fit result for the 2D screening length. From this we can conclude that for *high vacancy concentrations* as shown in Fig. 3, the screening potential is better described by the 2D surface screening potential model in Eq. (4).

4. Origin of the surface screening

At this stage we address, from where the surface screening charges arise. At first view, cleaved InP and GaAs(110) surfaces have no surface free charges, because no intrinsic surface states are within the fundamental band gap. Thus, on defect-free surfaces only bulk charge carriers can screen the localized vacancy charge. However, the vacancies themselves introduce defect states within the fundamental band gap, which are the origin of the positive charge of the vacancies. At high vacancy concentrations the individual vacancy defect states overlap and form a defect band as, e.g., calculated and illustrated by Qian *et al.*[46] The defect band labeled β in Ref. 46 is in the upper part of the band gap and it is empty for the $+1e$ charged vacancy and filled with one electron for each uncharged vacancy. At the vacancy concentrations used, the defect band can be partially filled, due to a statistical fraction of the vacancies in a neutral charge state. Thereby free electrons are introduced at the surface, which can screen the localized charges.

The population with electrons of the defect band is determined by the so-called charge-transfer level $E(+/0)$ from a $+1e$ to the neutral charge state. When the Fermi energy is at the charge-transfer level, half of the vacancies are in each charge state at any moment. For P vacancies on *p*-doped InP(110) surfaces the charge-transfer level from $+1e$ to 0 has been determined experimentally[47] to $E(+/0) = E_V + 0.75$ eV, with E_V being the valence-band maximum, in agreement with theory.[45] Since at a vacancy concentration of 1% ($= 4.1 \times 10^{12}$ cm^{-2}) the Fermi energy is at about ([$E_V + 0.65] \pm 0.03$) eV as measured by combined photoelectron spectroscopy and STM measurements (see Fig. 2 in Ref. 47), a sizable fraction of the vacancies will already be in the neutral charge state at any moment[48] due to the smearing out of the Fermi-Dirac distribution. The fraction p of uncharged vacancies can be estimated using $p = \exp[\{E_F - E(+/0)\}/kT]$. With the annealing temperature of 480 K and the above-mentioned energy levels, we obtain $p \approx 0.08^{+0.09}_{-0.05}$. Thus, the defect band will have a sufficient charge-carrier concentration available for screening and thereby the vacancies create the screening charges themselves.

Note, in Fig. 3 we fitted the interaction potential assuming a charge $q = +1e$ (i.e., $p = 0$) Since a fraction of vacancies is uncharged, the average vacancy charge[48] is somewhat smaller than $+1e$, i.e., $q = (1 - p)e$. Taking this into account in the fitting of the interaction energy data, yields slightly increased surface screening lengths R_{surf}. However, the values are subject to a rather large uncertainty, because the two fitting parameters, R_{surf} and p, are not independent [they are related in Eq. (5) through E_F, which depends of the carrier concentration]. Furthermore for a self-consistent iterative fit one needs to assume a value for the effective mass of the screening carriers. Due to the lack of any quantitative knowledge of the effective mass of the defect band, and the rather small effect on the screening length, we approximate the average vacancy charge in the following still by $+1e$. This will also simplify the comparison with smaller vacancy concentrations, where no defect band is populated.

Finally, we mentioned above that some vacancies appear uncharged in STM images due to a nearby dopant atom. Also these vacancies provide a charge carrier to the defect band. However, the dopant atom concentration close to the surface is rather low compared to the vacancy concentration. Thus this channel of free-electron provision is secondary and needs not to be taken into account separately.

C. The dependence of the screening length on the carrier concentration and vacancy concentration

Thus far we discussed a representative data for high vacancy concentrations only. In order to obtain a full overview, we measured a large number of data sets with strongly varying vacancy concentrations (0.049–1.29 %) and different doping concentrations (1×10^{18} and 3×10^{20} cm^{-3}) on two materials, InP and GaAs. From each data set we extracted the vacancy positions and the interaction energy. Each interaction energy is then fitted in an analogous manner as shown in Fig. 3. The resulting screening lengths are shown in Fig. 4 as a function of the effective carrier concentration at the surface. Filled symbols represent screening length obtained by fitting the screened Coulomb potential [Eq. (2)], whereas empty symbols represent fits by the surface screening potential in Eq. (4). The effective carrier concentration at the surface represents the concentration of free charge carrier at and near the surface available for bulklike screening of the vacancy charges in the absence of any surface screening.

The effective carrier concentration at the surface is determined by the bulk doping concentration and the density of surface vacancies. The vacancies induce a band bending at the surface, which leads to a reduced carrier concentration at the surface. The band bending can be calculated from the equilibrium between the surface charge density Q_{ss} (given by the number of surface vacancies times their charge) and the charge density Q_{sc} in the depletion layer compensating the surface charge density [see Eqs. (1) and (2) in Ref. 47]. With the band bending one can calculate the effective carrier concentration near the surface, using the dependence of the carrier concentration on the relative position of the Fermi energy with respect to the shifted valence-band edge at the surface [see Eq. (3) in Ref. 47]. For this calculation we used charge-transfer levels $E(+/0)$ for P and As vacancies on InP and GaAs(110) surfaces of $E_V + 0.75$ eV (Ref. 47) and $E_V + 0.49$ eV,[45] respectively. The thus obtained effective surface carrier concentration is used as *x* axis in Fig. 4.
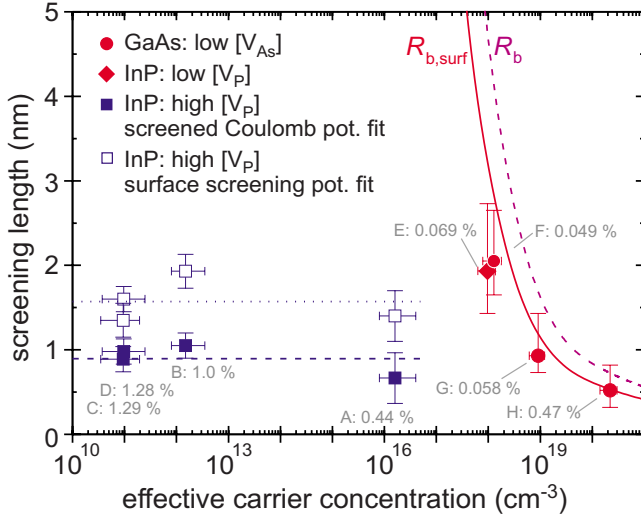
FIG. 4. (Color online) Screening length as determined from fits of the screened Coulomb potential [Eq. (2)] and the surface screening potential [Eq. (4)] versus the effective carrier concentration at the surface. For each data set (labeled A–H) the vacancy concentration is indicated. Two different screening mechanisms are visible. At low vacancy concentrations (high effective carrier concentrations, red circles and diamond) a bulk screening occurs, whereas at high vacancy concentrations a surface screening dominates (low carrier concentrations, blue squares). The transition between 2D and 3D screening arises from the formation of a partially filled defect band in the band gap.

Several observations are worth noting in Fig. 4. (i) If we compare only the screening length obtained from a screened Coulomb potential fit (blue filled squares), the values at low effective carrier concentrations are essentially constant over many orders of magnitude of effective carrier concentration. The same behavior is also found if we analyze the surface screening length obtained from a fit using the surface screening potential (blue empty squares).

(ii) At high effective carrier concentrations the screening length decreases strongly with increasing carrier concentration (red filled circles and diamond). (iii) The two different dependencies are connected to high and low vacancy concentrations, which are indicated near the blue and red symbols, respectively. The higher vacancy concentrations yield the constant screening length apparently independent of the effective carrier concentration. (iv) The data indicates a screening transition with increasing vacancy concentration.

First, we discuss the constant screening length at low effective carrier concentrations. These arise from the surface screening through carriers in the (vacancy) defect band as discussed above for Fig. 3. No dependence on the effective (bulklike) carrier concentration near the surface is observed (independent of the potential used for the fit), because the screening is completely dominated by surface screening. The bulk screening length would be extremely large (see dashed line labeled R_b in Fig. 4) at these low effective carrier concentrations and thus irrelevant. Since we observe this regime of surface screening only at high vacancy concentrations, we have very similar concentrations of surface screening charges in the defect band and thus similar screening lengths

according to Eq. (5). Note, although this regime is dominated by surface screening [Eq. (4)] and thus the correct screening lengths are given by Eq. (5) [open (blue) symbols], we also show the screening lengths obtained by fitting the screened Coulomb potential (filled blue symbols) for comparison with the bulk screening at high effective carrier concentrations (low vacancy concentration).

Second, the screening lengths obtained for samples with a low vacancy concentration (red filled circles and diamond at high carrier concentrations in Fig. 4) have a different behavior. This arises from the large separation between the vacancies, which does not allow the formation of the defect band at the surface. Thus, no surface screening charges are present and only bulk free charge carriers screen the localized charges of the vacancies. Therefore, the screening lengths should be described by Eq. (3), i.e., the dashed line marked R_b in Fig. 4. The dashed line illustrates that the bulk screening length R_b increases strongly with decreasing carrier concentration. The experimental data points also exhibit this trend. However, it appears that they are shifted somewhat. This shift suggests that one needs to consider in Eq. (3) that the dielectric constant ε_r is different at the surface [$\varepsilon_{surf} = (\varepsilon_r + 1)/2$]. Assuming therefore for the bulklike screening length at the surface $R_{b,surf} = R_b(\varepsilon_{surf})$ one obtains the solid red curve, which shows a much better agreement with the experimental data. This suggests that the screening length in Eq. (3) may need to be corrected for the surface dielectric constant.

Third, the screening of surface charges exhibits a dimensional transition with increasing vacancy concentration visible in Fig. 4 as different functional dependencies between the high vacancy (at low effective carrier concentration, blue symbols) and the low vacancy concentration (at high effective carrier concentrations, red symbols) data. This indicates that initially with only individual vacancies at the surface, screening is governed exclusively by bulk charge carriers. Once the vacancy concentration is increasing, such that a notable band bending occurs and thereby the defect band formed at high vacancy concentrations can be partially filled, then the screening is governed by the charge carriers in the partially filled defect band. Thus a 3D- to 2D-screening transition occurs with increasing vacancy concentration.

Note, the data point with the highest carrier concentration in Fig. 4 has also a relatively high vacancy concentration. However, due to the very large doping concentration, the band bending induced by the vacancies is insufficient to fill the defect band. Thus, no surface screening charges are present in that case and only screening by bulk charge carriers occurs.

Finally, a 3D- to 2D-screening transition can also occur in the bulk, when dopant atoms spontaneously reorder and form 2D accumulations with 2D charge carrier characteristics as reported in Ref. 49. In the present case the two dimensionality of the screening is due to the surface and independent of special ordering effects of bulk dopant.[49]

IV. DETERMINATION OF THE SCREENING LENGTH IN STM IMAGES

The screening length of charged surface defects has been determined in many occurrences directly from height profiles

in constant-current STM images.[9–14] This determination assumes that the apparent height change Δz in the STM image is given by $\Delta z = \Delta z_0 \times \exp(-r/R_s) + \mathrm{const.}$,[10] with R_s being the screening length of the assumed screened Coulomb potential. These data always yielded screening length in the order of 1 nm independent of the doping level of the material investigated. Another example is that the imaged screening potential around dopants is independent of the location in the vicinity of p-n junctions, i.e., the depletion zone or the homogeneously doped area.[50] In the light of the above presented observation of a purely 2D screening in the presence of a surface band, one can now reinterpret the data directly measured in STM images. In the presence of the scanning tunneling microscope tip in close proximity to the sample, the tip induces a *localized* band bending at the semiconductor surface.[16,17] This band bending has typically the extension of the tip's radius of curvature, i.e., in the range of 100 nm. In addition the tip can induce electron states within the band bending area.[16,17] If these energy levels are partially filled, they form a 2D surface charge-carrier system, which screens the localized defect charges. This screening follows in analogy to our high vacancy concentration regime the 2D surface screening regime, yielding rather short screening length defined by Eq. (5) in the range of 1–2 nm. This value is in good agreement with the typical screening lengths of 1 nm determined directly from the STM images. Note, that due to the different functional form of the 2D and 3D screening potentials, the resulting screening length differs roughly by a factor 2, i.e., for the same experimental potential one obtains a screening length twice as large when using a fit based on the surface screening potential versus the screened Coulomb potential (compare also with our screening lengths in Figs. 3 and 4). Therefore, the 1 nm screening length determined from height profiles using the exponential screening factor of the screened Coulomb potential, corresponds rather to a 2D screening length of roughly 2 nm.

V. CONCLUSIONS

We investigated quantitatively the screening of localized charges on p-doped InP and GaAs(110) surfaces. As model system for probing the screening potentials we use the electrostatic interaction of positively charged anion vacancies, formed thermally on the surfaces. In order to avoid the influence of the tips' electric field on the screening potential, the interaction potentials were extracted from the statistical spatial distribution of the vacancies. Two screening regimes are identified: at low vacancy concentrations screening is governed by bulk charge carriers. The screening length is found to increase strongly with decreasing carrier concentration in agreement with the expectation for three-dimensional bulk screening. The comparison of the experimental data and the theoretical values of the bulk screening length suggest that the best description is achieved, when evaluating the bulk screening length using for the surface dielectric constant the average of the semiconductor and the vacuum. At high vacancy concentrations a two-dimensional surface screening potential with small screening lengths is observed, independent of the effective bulk carrier concentration. The screening is governed by charge carriers in a two-dimensional surface vacancy defect band, which is partially populated by the large surface band bending induced by the high vacancy concentration. Thus, the screening undergoes a three- to two-dimensional transition with increasing vacancy concentration.

*Present address: OSRAM Opto-Semiconductors GmbH, Leibnizstrasse 4, 93055 Regensburg, Germany.

†p.ebert@fz-juelich.de

[1] R. B. Dingle, Philos. Mag. **46**, 831 (1955).

[2] R. Mansfield, Proc. Phys. Soc. London, Sect. B **69**, 76 (1956).

[3] P. Csavinszky, Philos. Mag. **10**, 231 (1964).

[4] F. Stern and W. E. Howard, Phys. Rev. **163**, 816 (1967).

[5] F. Stern, Phys. Rev. Lett. **18**, 546 (1967).

[6] M. Krčmar, W. M. Saslow, and M. B. Weimer, Phys. Rev. B **61**, 13821 (2000); **62**, 7657 (2000) (erratum); **64**, 199901 (2001) (erratum).

[7] M. R. Krčmar and W. M. Saslow, Phys. Rev. B **66**, 235310 (2002).

[8] R. Slavchov, T. Ivanov, and B. Radoev, J. Phys.: Condens. Matter **18**, 5873 (2006).

[9] J. A. Stroscio, R. M. Feenstra, and A. P. Fein, Phys. Rev. Lett. **58**, 1668 (1987).

[10] R. J. Hamers, J. Vac. Sci. Technol. B **6**, 1462 (1988).

[11] G. Cox, K. H. Graf, D. Szynka, U. Poppe, and K. Urban, Vacuum **41**, 591 (1990).

[12] Ph. Ebert and K. Urban, Ultramicroscopy **49**, 344 (1993).

[13] G. Lengel, R. Wilkens, G. Brown, and M. Weimer, J. Vac. Sci. Technol. B **11**, 1472 (1993).

[14] Ph. Ebert, X. Chen, M. Heinrich, M. Simon, K. Urban, and M. G. Lagally, Phys. Rev. Lett. **76**, 2089 (1996).

[15] R. M. Feenstra and J. A. Stroscio, J. Vac. Sci. Technol. B **5**, 923 (1987).

[16] R. Dombrowski, Chr. Steinebach, Chr. Wittneven, M. Morgenstern, and R. Wiesendanger, Phys. Rev. B **59**, 8043 (1999).

[17] M. Wenderoth, M. A. Rosentreter, A. J. Heinrich, M. A. Schneider, and R. G. Ulbrich, Europhys. Lett. **45**, 579 (1999).

[18] N. D. Jäger, M. Marso, M. Salmeron, E. R. Weber, K. Urban, and Ph. Ebert, Phys. Rev. B **67**, 165307 (2003).

[19] K. Kobayashi, S. Kurokawa, and A. Sakai, Jpn. J. Appl. Phys., Part 1 **43**, 4571 (2004).

[20] I. Nevo, S. Aloni, S. R. Cohen, and G. Hasse, J. Chem. Phys. **123**, 064705 (2005).

[21] S. Yoshida, Y. Kanitani, O. Takeuchi, and H. Shigekawa, Appl.

Phys. Lett. **92**, 102105 (2008).

[22] Ph. Ebert, M. Heinrich, M. Simon, K. Urban, and M. G. Lagally, Phys. Rev. B **51**, 9696 (1995); M. Heinrich, Ph. Ebert, M. Simon, K. Urban, and M. G. Lagally, J. Vac. Sci. Technol. A **13**, 1714 (1995).

[23] C. Domke, M. Heinrich, Ph. Ebert, and K. Urban, J. Vac. Sci. Technol. B **16**, 2825 (1998).

[24] Ph. Ebert, G. Cox, U. Poppe, and K. Urban, Ultramicroscopy **42-44**, 871 (1992).

[25] S. Landrock, Y. Jiang, K. H. Wu, E. G. Wang, K. Urban, and Ph. Ebert, Appl. Phys. Lett. **95**, 072107 (2009).

[26] T. L. Hill, *Statistical Mechanics* (McGraw-Hill, New York, 1956).

[27] A. Laubsch, Diploma thesis, RWTH Aachen University, 2005.

[28] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 2nd ed. (Academic, London, 1986), Chap. 5.

[29] Ph. Ebert, Tianjiao Zhang, F. Kluge, M. Simon, Zhenyu Zhang, and K. Urban, Phys. Rev. Lett. **83**, 757 (1999).

[30] Ph. Ebert, Surf. Sci. Rep. **33**, 121 (1999); Curr. Opin. Solid State Mater. Sci. **5**, 211 (2001).

[31] Ph. Ebert, M. G. Lagally, and K. Urban, Phys. Rev. Lett. **70**, 1437 (1993).

[32] G. Lengel, J. Harper, and M. Weimer, Phys. Rev. Lett. **76**, 4725 (1996).

[33] G. Lengel, M. Weimer, J. Gryko, and R. E. Allen, J. Vac. Sci. Technol. B **13**, 1144 (1995).

[34] Ph. Ebert and K. Urban, Surf. Sci. **287-288**, 891 (1993).

[35] R. M. Feenstra, J. A. Stroscio, J. Tersoff, and A. P. Fein, Phys. Rev. Lett. **58**, 1192 (1987); Ph. Ebert, B. Engels, P. Richard, K. Schroeder, S. Blügel, C. Domke, M. Heinrich, and K. Urban, *ibid.* **77**, 2997 (1996); B. Engels, P. Richard, K. Schroeder, S. Blügel, Ph. Ebert, and K. Urban, Phys. Rev. B **58**, 7799 (1998); Ph. Ebert, G. Cox, U. Poppe, and K. Urban, Surf. Sci. **271**, 587 (1992); N. D. Jäger, E. R. Weber, K. Urban, and Ph. Ebert, Phys. Rev. B **67**, 165327 (2003).

[36] K. J. Chao, A. R. Smith, and C. K. Shih, Phys. Rev. B **53**, 6935 (1996); Ph. Ebert, M. Heinrich, M. Simon, C. Domke, K. Urban, C. K. Shih, M. B. Webb, and M. G. Lagally, *ibid.* **53**, 4580 (1996).

[37] M. B. Johnson, O. Albrektsen, R. M. Feenstra, and H. W. M. Salemink, Appl. Phys. Lett. **63**, 2923 (1993); **64**, 1454 (1994) (erratum).

[38] G. Lengel, R. Wilkins, G. Brown, M. Weimer, J. Gryko, and R. E. Allen, Phys. Rev. Lett. **72**, 836 (1994).

[39] Ph. Ebert, K. Urban, and M. G. Lagally, Phys. Rev. Lett. **72**, 840 (1994).

[40] Equation (2) differs slightly from Eq. (35) in Ref. 6, because we used a slightly different but equally well-suited approximation, which yields the well-known relation that the dielectric constant at an interface is the average of the two media, i.e., in our case $\varepsilon_{surf}=(\varepsilon_r+1)/2$ [Eq. (32) in Ref. 4; Eq. (2.11) in T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982)]. Reference 6 uses effectively $\varepsilon_{surf}=\varepsilon_r/2$.

[41] Equation (B19) in Ref. 4 in the limit $d \rightarrow 0$ and Eq. (16) in Ref. 7.

[42] $F_k(\eta)=1/k! \times \int_0^\infty x^k/(e^{x-\eta}+1)dx$ is, in principle, not defined for $k=-1$ due to a singularity in the Γ function defining $k!$. But the integral is also divergent and the divergencies cancel out each other. Therefore we defined $F_{-1}(\eta)=\lim_{k\rightarrow-1} F_k(\eta)$. With the relation $F'_k(\eta)=F_{k-1}(\eta)$, this yields $F_{-1}(\eta)=1/(1+e^{-\eta})$.

[43] In order to fit the surface screening potential [Eq. (4)] to the experimental data, we used a series expansion for $H_0(u)$-$N_0(u)$.

[44] G. Schwarz, Ph.D. thesis, Technical University Berlin, 2002.

[45] M. Hedström, A. Schindlemayr, G. Schwarz, and M. Scheffler, Phys. Rev. Lett. **97**, 226401 (2006).

[46] M. C. Qian, M. Göthelid, B. Johansson, and S. Mirbt, Phys. Rev. B **66**, 155326 (2002).

[47] Ph. Ebert, K. Urban, L. Aballe, C. H. Chen, K. Horn, G. Schwarz, J. Neugebauer, and M. Scheffler, Phys. Rev. Lett. **84**, 5816 (2000).

[48] One effectively observes a time average charge state of the vacancy in the STM images, since the vacancies will change their charge state statistically at much higher frequencies than the STM can detect.

[49] S. Landrock, K. Urban, and Ph. Ebert, Phys. Rev. Lett. **96**, 076101 (2006).

[50] N. D. Jäger, K. Urban, E. R. Weber, and Ph. Ebert, Phys. Rev. B **65**, 235302 (2002); Appl. Phys. Lett. **82**, 2700 (2003).