PART ONE

1

Introduction

E. F. SCHUBERT

This book is devoted to semiconductors containing very thin, highly doped layers. Assuming that doping impurities are distributed uniformly within the two-dimensional layer, the doping profile depends on only one spatial dimension. Frequently, such doping distributions are described by the delta (δ-) function. This chapter will introduce the reader to the field of δ-doping. Section 1.1 will address the questions: What are δ-doped semiconductors? How are they fabricated? Why are they made? In Section 1.2, a historical sketch of this subject will be provided. The electronic structure of δ-doped semiconductors will be derived in Section 1.3. Finally, the implications of δ-doping on state-of-the-art microelectronic devices will be discussed in Section 1.4.

1.1 Delta doping: what is it and how is it done?

Delta doping is motivated by the need for well-defined, narrow doping profiles in semiconductors. Doping distributions with high peak concentrations and narrow distribution widths are advantageous for many device applications. The narrowest one-dimensional (1D) doping profile is achieved if doping atoms are confined to a single atomic layer in the host semiconductor. Such a situation is illustrated schematically in Fig. 1.1, which shows a semiconductor substrate, an epitaxial layer, and a dopant sheet embedded in the epitaxial layer. All impurities, shown as dark circles in the inset to Fig. 1.1, occupy substitutional lattice sites in the host semiconductor. In group-IV semiconductors, all lattice sites are identical. In III–V semiconductors, we have the distinct cation and anion lattice sites. Therefore, the doping plane shown in Fig. 1.1 may be either an anion or a cation sublattice plane.

The two parameters characterizing a δ-doping profile are the location of the dopant sheet and the density of doping atoms in the sheet. Assuming that the wafer surface is in the xy-plane of a cartesian coordinate system, that the z-coordinate is measured from the surface to the bulk, and that the dopants are located in the plane at $z = z_d$, then the doping

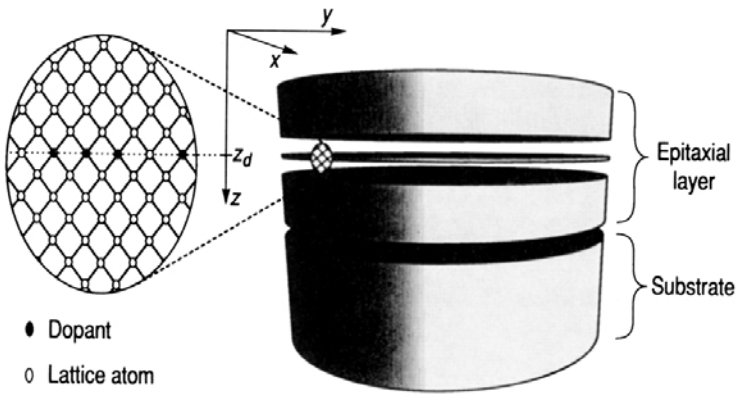

Fig. 1.1. Schematic illustration of a semiconductor substrate and an epitaxial film containing a δ-doping layer. Also shown is a schematic lattice with the impurity atoms being confined to a single atomic plane.

profile is given by

$$N(z) = N^{2D}\delta(z - z_d) \tag{1.1}$$

where the two-dimensional (2D) density, denoted by N^{2D}, is the number of doping atoms in the doping plane per cm^2. Equation (1.1) states that the doping concentration is zero for all locations, except for $z = z_d$. Integration of Eqn. (1.1) yields $\int_{-\infty}^{+\infty} N(z)\,dz = N^{2D}$, that is, the 2D doping density.

The growth of δ-doped semiconductors will be discussed for different epitaxial growth techniques in Chapter 3. Here, we summarize briefly the basic procedure common to all growth techniques. Consider a semiconductor during epitaxial growth. In the *first step* of δ-doping, the epitaxial growth of the semiconductor is suspended, which can be achieved by closing the effusion cells in molecular-beam epitaxy or by valving off the semiconductor precursor flow in vapor-phase epitaxy. We assume that the non-growing semiconductor surface will be atomically flat with no atomic terrace steps on the surface. In the *second step*, the surface is exposed to either a flux of elemental doping atoms or to the flux of a doping precursor. For efficient incorporation, it is desirable that the doping atoms have a sticking probability near unity, or, for the case of a precursor, that the decomposition probability of the precursor is near unity. If the doping atoms do not re-evaporate, then they can form bonds with host semiconductor atoms. Impurity atoms are *chemisorbed* at the semiconductor surface. Ideally, all doping atoms occupy substitutional lattice sites on the semiconductor surface. In the final *third step* of δ-doping, the deposition of dopants is terminated and the epitaxial growth of the host semiconductor is resumed. In the absence of impurity redistribution effects such as diffusion, drift, and segregation, the doping atoms are then confined to one atomic layer of the semiconductor.

However, if the semiconductor surface on which the impurities are deposited is rough and contains steps or terraces, or if thermally stimulated redistribution of impurities does occur during subsequent growth, then the impurities cannot be confined to a single atomic

layer. Further details of the epitaxial growth of δ-doping layers using different techniques will be discussed in Part 3.

The density of doping atoms in the doping monolayer can be determined by the dopant flux, which may be a flux of doping atoms or of precursor molecules, and the time that the semiconductor surface is exposed to the dopant flux. Thus, for a given dopant flux, the time t, during which the semiconductor is exposed to the dopant flux, can be varied to achieve the desired 2D density of incorporated dopants. Frequently, the bulk concentration (3D concentration) of a dopant is known for a given set of experimental conditions such as a given growth rate, substrate temperature, effusion cell temperatures, gas flows, etc. If this bulk concentration is denoted by N, then the 2D dopant density obtained during δ-doping is given by

$$N^{2D} = N v_g t \tag{1.2}$$

where N is the bulk concentration (per cm^3) obtained at a given growth rate v_g, and t is the time for which the growth is suspended. Equation (1.2) thus allows one to select the time t required to obtain a doping density N^{2D} provided that the bulk concentration N is known for a given set of growth conditions. Equation (1.2) is valid if (i) dopants deposited on the non-growing surface have the same incorporation probability as dopants incorporated during a continuous growth mode and if (ii) the growth rate of the host semiconductor is not influenced by the doping incorporation. Both conditions are usually met by epitaxial growth systems.

The effective 3D concentration of dopants in the doping sheet of a δ-doped semiconductor can be inferred from the 2D density. The electrical activity of some dopants decreases at high concentrations. Other doping impurities diffuse strongly at high concentrations. It is therefore desirable to estimate the effective 3D concentration of dopants in a 2D doping sheet. In order to calculate the effective 3D concentration, we consider a sheet of dopants with density N^{2D}. The mean distance between the doping atoms is $(N^{2D})^{-1/2}$. Now consider a homogeneously doped semiconductor with *the same mean distance between the dopant atoms*. This semiconductor would have an *equivalent 3D doping concentration* of

$$N^{3D} = (N^{2D})^{3/2}. \tag{1.3}$$

The quantity N^{3D} is the physically relevant 'bulk' concentration of a δ-doped semiconductor. For example, a δ-doping concentration of $N^{2D} = 10^{12}\ cm^{-2}$ corresponds to an equivalent 3D concentration of $N^{3D} = 10^{18}\ cm^{-3}$, as calculated from Eqn. (1.3).

In actual δ-doped semiconductors, dopants may not be confined to a single atomic layer but may be distributed over more than a single layer. Surface roughness and other processes such as diffusion, drift, and segregation may contribute to the doping redistribution. In order to quantify the spread of dopants in δ-doped semiconductors, we first consider, as the simplest approximation, a top-hat distribution, which is illustrated schematically in Fig. 1.2. The width of the doping distribution, Δz_d, is smaller than the lattice constant, a_0. However, for sufficiently strong doping redistribution processes, the distribution width, Δz_d, may become larger than the lattice constant, a_0. For all well-behaved doping

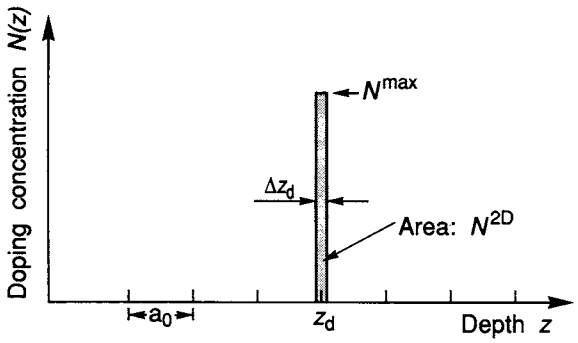

Fig. 1.2. Doping profile of a doping spike located at $z = z_d$ with 2D density N^{2D}, maximum concentration N^{max}, and full-width at half-maximum of Δz_d.

profiles such as a top-hat, a gaussian, or a triangular doping distribution, the maximum concentration and the profile width are related by

$$N^{2D} \approx N^{max} \Delta z_d. \tag{1.4}$$

Considering redistribution processes, the question arises as to whether doping profiles with $\Delta z_d \gtrsim a_0$ can be considered to be δ-function-like. To answer this question, the actual distribution width Δz_d must be compared to other relevant length scales such as the screening length, free-carrier diffusion length, depletion length, free-carrier wavelength, etc. More specifically, the distribution width Δz_d must be smaller than the *shortest* of these length scales. The shortest of the above-mentioned length scales is, in most cases, the free-carrier de Broglie wavelength. The free-carrier de Broglie wavelength decreases with increasing effective mass and the carrier kinetic energy. In practice, the free-carrier de Broglie wavelength is longer than 25 Å, even for carriers with large effective mass and high kinetic energy. We can therefore consider all doping profiles with a distribution width of $\Delta z_d \leqslant 25$ Å to be δ-function-like.

As stated above, δ-doped semiconductors are predominantly fabricated by epitaxial growth. Other doping techniques, such as doping by diffusion and by ion implantation, produce inherently broader doping distributions. However, ion implantation can produce spatially well-confined doping profiles if the implantation depth is very shallow. As an approximate rule, the projected straggle of implanted ions equals the implantation depth. The redistribution of doping ions during postimplantation annealing can be minimized by rapid thermal annealing. Therefore, shallow implantation of impurities with implantation depths of less than 100 Å can produce near-δ-function-like doping profiles.

1.2 Historical review

During the δ-doping process, the non-growing crystal surface is exposed to a flux of doping atoms or doping precursor molecules. These atoms or molecules undergo chemisorption

on the non-growing crystal surface. The chemisorption is a desirable effect. However, the chemisorption of impurities on a non-growing semiconductor surface may also be an undesirable process, namely if it occurs before the epitaxial growth is initiated. Before a wafer is introduced into the growth chamber, it is exposed to the ambient, usually clean air. After the introduction of a wafer into the growth system, the wafer is exposed to the ambient of the system, usually a carrier gas or vacuum. The exposure time for the wafer can be several hours, until the actual epitaxial growth is initiated. The incoporation of impurities at the substrate–epilayer interface was first reported by DiLorenzo (1971), who found a high concentration of Si, C, Mg, Fe, Cr, and other impurities at the substrate–epilayer interface. These impurities were incorporated during the substrate cleaning procedure, as well as during the exposure of the wafers to the carrier gas of the vapor-phase growth system. Considering the state-of-the-art of substrate cleaning in 1971, most impurities were probably incorporated during substrate cleaning.

The incorporation of impurities at the substrate–epilayer interface was confirmed by Bass (1979), who showed that adsorption (chemisorption and physisorption) of impurities on the substrate surface can occur in the reaction chamber before epitaxial growth is initiated. During GaAs growth by vapor-phase epitaxy (VPE), hydrogen sulfide (H_2S), silane (SiH_4), and germane (GeH_4) were used as doping precursors (Bass, 1979). It was general practice to initiate the doping precursor flow before the actual GaAs growth was started. This was done to eliminate any transient effects occurring in the H_2S doping line predominantly, and, to a smaller extent, in the silane and germane doping lines. When the purging procedure was done with silane, a pronounced doping spike occurred at the substrate–epitaxial layer interface. This indicated that silicon was strongly chemisorbed on the GaAs substrate surface. Doping profile measurements by the capacitance–voltage (CV) technique indeed revealed the occurrence of a clear doping spike at the substrate–epilayer interface. The measured full-width at half-maximum of the Si doping spikes exceeded 200 Å and the doping profiles can, therefore, not be considered as δ-function-like distributions in the strict sense defined above. Nevertheless, Bass's report (1979) was the first report of doping incorporation on a non-growing semiconductor crystal.

Realizing the possibility of narrow doping distributions, Bass (1979) proposed that the strong surface adsorption of silicon could be used to produce sharp doping spikes which could be used advantageously in IMPATT diodes. A strong adsorption of Si on the GaAs surface was also reported for epitaxial growth by molecular-beam epitaxy (MBE) (Wood *et al.*, 1980). However, the profile widths were >300 Å, that is, broader than those reported by Bass. Wood *et al.* (1980) proposed that the doping technique could be used to synthesize complex free-carrier profiles such as linearly ramped profiles.

The first truly δ-doped semiconductor structure with clear evidence for a narrow doping profile was reported by Schubert *et al.* (1984) and Schubert and Ploog (1985). Capacitance–voltage (CV) measurements on MBE grown GaAs samples revealed a half-width at half-maximum of 20 Å. Assuming that the profiles are symmetric, a full profile width of 40 Å can be deduced from the CV measurements. This profile width is at least five times narrower than the profile widths reported by Bass (1979) and Wood *et al.* (1980) and was

thus the first truly δ-doped semiconductor. In the same publication, Schubert and Ploog (1985) reported the first field-effect transistor (FET) using δ-doped GaAs. The authors showed that the δ-doped FET has a narrow free-carrier and dopant distribution and a large gate-breakdown voltage. Such FETs also exhibit reduced short-channel effects due to the narrow distribution of dopants.

The δ-doping technique has been used by Schubert and coworkers in a number of MBE-grown semiconductor devices. Examples are the homojunction FET (Schubert and Ploog, 1985), the high-mobility heterojunction FET (Schubert *et al.*, 1987), light-emitting diodes (LEDs) (Schubert *et al.*, 1985b), and lasers (Schubert *et al.*, 1985a, 1989a). When applied to heterojunction FETs, the δ-doping technique yields the highest free-carrier concentrations. It has been shown that δ-doping distributions are the optimum doping profiles for modulation-doped FETs. For LEDs and lasers, δ-doping has been used in the active region of the structure in which electron–hole recombination occurs. The active region in these optical devices consists of a doping superlattice which allows one to lower the bandgap energy of the active region. Emission wavelengths exceeding 950 nm for a GaAs host have been achieved at room temperature. In addition, lasing has been achieved in δ-doped superlattices (Schubert *et al.*, 1985a and 1989a). Constant lasing wavelengths, as well as tunable wavelengths, have been achieved with δ-doped superlattice lasers.

Delta-function-like doping distributions minimize potential fluctuations originating from random doping atom distributions. Consider a doped slab of concentration N and thickness Δz_d. Assume that the doping atoms are randomly distributed within this slab. Thus, potential fluctuations due to the random dopant distribution will be felt in the vicinity of the slab. One can show that such potential fluctuations, occurring at any point outside the doping slab, are minimized if the thickness of the doped region approaches zero ($\Delta z_d \to 0$), that is, for the δ-doped case (Schubert *et al.*, 1988a). Many characteristics of semiconductor structures are related to potential fluctuations, for example the mobility of free carriers or the luminescence linewidth of doped structures. These characteristics can be optimized by using the δ-doping technique. The mobility of selectivity doped heterostructures is maximized by δ-doping (Schubert *et al.*, 1989b). Free carriers in selectively doped heterostructures are scattered by potential fluctuations caused by the random distribution of remote dopants. Minimizing these fluctuations results in the minimization of scattering by remote ionized dopants. Maximum experimental mobilities have indeed been achieved in δ-doped heterostructures as opposed to homogeneously doped heterostructures (Schubert *et al.*, 1989b).

The minimization of potential fluctuations in δ-doped structures has resulted in a significant improvement of the optical properties of doping superlattices. Quantum-confined interband transitions have been observed in absorption measurements (Schubert *et al.*, 1988b) and luminescence emission experiments (Schubert *et al.*, 1989c). The measured peak energies have been assigned to theoretical transition energies and good agreement has been found between measured and calculated transition energies. Up to the present time, quantum-confined interband transitions have not been reported for homogeneously doped doping superlattices.

The assessment of the spatial distribution of doping atoms in δ-doped structures has been the subject of intense interest, since it is the condition *sine qua non* for δ-doped structures. The early work of Bass (1979) and Wood *et al.* (1980) indicated distribution widths >200 Å. It is likely that the broad profile widths were indeed caused by broad doping distributions rather than a low spatial resolution of the measurement technique. The clear signature of doping redistribution effects of Si in $Al_xGa_{1-x}As$ has also been reported by Lee *et al.* (1985), who concluded that Si δ-function-like profile widths exceed 100 Å.

The spatial localization of doping atoms in δ-doped semiconductors has been investigated using a number of characterization techniques including Auger electron spectroscopy, capacitance–voltage (CV) profiling, secondary ion mass spectrometry (SIMS), and transmission electron microscopy (TEM). The first study on the spatial localization of dopants revealed a 20 Å half-width, half-maximum CV-profile width and thus indicated very good confinement of dopants (Schubert and Ploog, 1985). In order to further study the range of temperatures suitable for the growth of Si δ-doped GaAs, localization was studied as a function of doping density and growth temperature (Schubert *et al.*, 1988c and 1990a). These studies revealed that the CV profile widths are limited by the resolution of the CV technique. Dopants were concluded to be confined to within 15 Å, that is, approximately three lattice constants.

Secondary ion mass spectrometry measurements were employed extensively in order to study the distribution of dopants in δ-doped semiconductors (Zeindl *et al.*, 1987; Beall *et al.*, 1988; Clegg and Beall, 1989; Lanzillotto *et al.*, 1989; Schubert *et al.*, 1990b). The SIMS technique has the highest resolution for shallow doping layers buried 200–500 Å below the semiconductor surface. Low sputtering energies of $\leqslant 3$ keV are advantageously used to achieve high resolution. Studies by different groups have revealed that the widths of the measured doping profiles are limited by the resolution of the SIMS technique. The narrowest profiles display a full-width at half-maximum of only 29 Å (Schubert *et al.*, 1990b). Spatial local-ization of dopants within a layer of thickness $\leqslant 15$ Å was concluded from the SIMS measurements.

The spatial localization and redistribution of Si in GaAs has also been studied by Webb (1989) using Auger electron spectroscopy. The study revealed that segregation of Si towards the GaAs surface during growth at 520 °C was immeasurably small. The Auger technique was shown to have an excellent resolution of approximately 5 Å.

Transmission electron microscopy (TEM) has been employed in the study of spatial localization of dopants in highly δ-doped semiconductors. Zeindl *et al.* (1987) showed that an Sb doping layer of density 5×10^{13} cm^{-2} can be clearly observed as a dark line in the bright-field lattice image of a Si specimen with a $\langle 011 \rangle$ oriented cross-section. The TEM measurements were made under conditions in which strain contrast was weak. The dark line was therefore assumed to be due to scattering of the electrons by heavy Sb atoms. The TEM micrograph indicated that the Sb doping atoms were confined to a layer of approximate thickness 15 Å. Thus, all techniques used to assess the doping distribution in δ-doped semiconductors, namely Auger, CV, SIMS, and TEM, revealed resolution-limited profiles for samples grown under optimized conditions.

The combination of high-resolution assessment techniques and δ-function-like doping profiles offers a sensitive method of measuring the redistribution and diffusion constant of dopants during high-temperature growth and postgrowth annealing. The redistribution of dopants during high-temperature growth can easily be detected by this method (Beall *et al.*, 1989; Schubert *et al.*, 1989). Postgrowth annealing of δ-doped samples for different times and/or temperatures allows one to determine the diffusion constant of dopants as a function of temperature (Schubert *et al.*, 1988d; Beall *et al.*, 1990). This method provides very high sensitivity as well as excellent accuracy for the assessment of diffusion constants.

The growth of δ-doped III–V semiconductors is achieved via the deposition of the doping element or of the doping precursor on the non-growing crystal surface at growth temperature. If this procedure is carried out during Si epitaxial growth, excessive segregation of impurities such as B or Sb results. In order to avoid this problem, Zeindl *et al.* (1987) first evaporated Sb dopants to the desired submonolayer coverage and then covered the doping layer by amorphous Si, all at room temperature. After amorphous growth, the Si wafer was heated to 700 °C to recrystallize the amorphous layer. This procedure is commonly referred to as solid-phase epitaxy. An alternative method of obtaining δ-function-like doping profiles in Si by MBE was demonstrated by Gossmann *et al.* (1990) and Gossman and Schubert (1993), who found that the Si growth temperatures can be reduced to 325 °C where the segregation of Sb and other doping elements is insignificantly small. The authors further showed that the quality of the films as assessed by Hall mobility measurements did not degrade due to the low-temperature growth.

It is commonly assumed that the doping atoms are randomly distributed within the doping plane. Potential fluctuations arise due to the random impurity distribution, which will scatter free carriers and reduce their drift mobility. If it were possible to place the doping atoms in an ordered, non-random arrangement, the carrier mobility would increase. Levi *et al.* (1989) proposed that, under suitable growth conditions, high-density δ-doping may result in partial ordering of the doping atoms. A reduction in the elastic scattering rate by up to a factor of four was predicted. To date, experimental evidence for ordered doping distributions has not been reported. However, Headrick *et al.* (1990 and 1991) reported on δ-doping with one full monolayer coverage of B in Si. The free-carrier mobilities were comparable to highly B-doped bulk Si.

Delta doping can reduce the adverse effects of persistent photoconductivity (PPC) and free-carrier freeze-out which have been found in n-type $Al_xGa_{1-x}As$. These effects have deleterious consequences for many semiconductor devices including microwave transistors. The physical origin of PPC and free-carrier freeze-out is a deep donor called the DX-center. Etienne and Thierry-Mieg (1988) found that the DX-center concentration is strongly reduced in Si δ-doped $Al_{0.3}Ga_{0.7}As$. The reduction is most pronounced at high-Si sheet densities.

Delta-function-like doping profiles were also grown by flow-rate modulation epitaxy (FME) which uses the same precursors as organo-metallic vapor-phase epitaxy (OMVPE). However, FME can grow at typically 100 °C lower temperatures as compared with conventional OMVPE. Sharp Si doping profiles with Si densities of 1×10^{13} cm^{-2} in GaAs grown at 550 °C have been demonstrated (Kobayashi *et al.*, 1986).

1.3 Basic theory of electronic structure

The free-carrier distribution in semiconductors depends on the distribution of ionized impurities. The free-carrier distribution is instrumental for many properties of semiconductors, including recombination and transport properties. In semiconductors with 'smooth' changes of doping concentration, the free-carrier profile follows the doping profile with good approximation. However, in δ-doped semiconductors which exhibit large doping concentration variations over short distances, the free-carrier distribution is spread out much further than the doping distribution. Here, we first discuss the mathematical representation of doping profiles in δ-doped semiconductors and subsequently discuss the free-carrier distribution in the classical as well as the quantum-mechanical picture.

The Dirac δ-function, $\delta(z)$, is a mathematical distribution whose width is infinitely narrow. The magnitude of $\delta(z)$ is zero for $z \neq 0$ and infinite for $z = 0$. The integral of the function, $\int_{-\infty}^{\infty} \delta(z)\,dz = 1$, has a value of unity. These are well-known properties of the δ-function. Employment of the δ-function for semiconductor doping profiles implies that the width of the one-dimensional profile is much narrower than *all other relevant length scales*. These relevant length scales include the screening length, the free-carrier diffusion length, and the de Broglie wavelength of a quantized carrier system.

Strickly speaking, the employment of the δ-function for any real physical quantity is problematic. As mentioned above, the δ-function approaches infinity, that is, $\delta(z \to 0) = \infty$, a value which cannot be assumed by any real physical quantity. As will be seen below, however, in the main the *integral* properties of the δ-function enter into calculations of the characteristics of δ-doped semiconductors. For example, the band potential of a semiconductor is obtained by *integration* of the δ-function-like charge distribution using the Poisson equation. Thus, the singularity of the δ-function at $z = 0$ does *not* impose a problem.

Alternatively, the gaussian function can be used to describe very narrow 1D doping distributions. The gaussian function is a continuous function and does not have a singularity. It is well known that the gaussian distribution function is characterized by the *standard deviation*. If the standard deviation approaches zero, then the gaussian function becomes identical to the δ-function. The simplicity and practicality of the δ-function make it the preferred choice for representing narrow doping profiles.

Next we will calculate several elementary quantities of δ-doped semiconductors. As an example, we consider an n-type semiconductor with a 2D donor density N_D^{2D}. We assume that the donor distribution depends on the spatial coordinate z only:

$$N_D(z) = N_D^{2D}\delta(z - z_d) \tag{1.5}$$

where z_d is the location of the plane of doping atoms. The doping profile of Eqn. (1.5) is illustrated schematically in Fig. 1.3(a). Assuming that all donors are ionized, the electrostatic potential created by the donor ions can be calculated using Poisson's equation.