Maximum concentration of impurities in semiconductors

E. F. Schubert, G. H. Gilmer, R. F. Kopf, and H. S. Luftman
AT&T Bell Laboratories, Murray Hill, New Jersey 07974
(Received 16 March 1992; revised manuscript received 22 July 1992)

The electronic properties of impurities are shown to limit the attainable maximum impurity concentration in semiconductors. The repulsive interaction between impurities due to (i) the Coulombic charge of ionized impurities and (ii) the increase of electronic energy at high doping concentrations can result in impurity segregation effects that limit the maximum achievable doping concentration. As an example, we consider the doping properties of Be, C, and Zn in GaAs. The characteristics of these impurities at extremely high concentrations agree with Monte Carlo simulations and molecular-dynamics calculations of impurity incorporation during crystal growth, which take into account the repulsive interaction between impurities.

Extremely high impurity concentrations in semiconductors become increasingly important in small semiconductor structures. Such structures have spatial dimensions on the order of the de Broglie electron or hole wavelength, i.e., in the 10–1000-Å scale. Consider, for example, a semiconductor cube with a volume of $100\times100\times100$ Å^3. Consider further that a statistically meaningful number of impurity atoms, e.g., 100 or 1000, reside within this quantum box. Such a number of impurities is required in order to keep statistical fluctuations sufficiently small. From this example it is evident that impurity concentrations on the order of 10^{20}–10^{21} cm^{-3} are required in semiconductor quantum structures. The requirement of extremely high impurity concentrations necessitates the understanding of the fundamental limits of impurity concentrations in semiconductors.

Several theories and models are available to explain the maximum concentration of impurities in semiconductors. Most prominent is the *solid-solubility limit,*[1] which was extensively applied to impurities in Si. However, the solid-solubility limit applies to thermal-equilibrium conditions which are usually not fulfilled under low-temperature growth conditions. Several groups showed that impurity concentrations can be achieved that exceed the so-called solid-solubility limit by several orders of magnitude.[2] Another theory to explain the maximum carrier concentration in compound semiconductors involves deep levels. Theis, Mooney, and Wright[3] proposed that the saturation of the carrier concentration in Si-doped GaAs is due to a Si-related level, the DX center, which becomes occupied at sufficiently high doping concentrations. Another interpretation of the maximum-impurity-concentration limit is based on the dependence of point-defect concentrations on the doping concentration.[4] However, this concentration-limiting mechanism is applicable only if the induced point defects exactly compensate the intentional doping. Nevertheless, these models are of very limited validity and do not provide a general explanation for the maximum achievable impurity concentration in semiconductors.

In this paper we propose a new model for the maximum impurity concentration in semiconductors. The model is based on the repulsive interaction between impurities due to (i) repulsive Coulombic interactions and (ii) the increase in electronic energy at high doping concentrations. The increase in energy leads to redistribution effects during epitaxial growth that effectively limit the maximum doping concentration. We will first present experimental results, then discuss the principle of the model, and finally present the results of computer calculations on impurity redistribution.

In order to study the redistribution of impurities at very high doping concentrations, Be-doped epitaxial GaAs layers were grown by molecular-beam epitaxy on (001)-oriented GaAs substrates. The samples consist of a 1-μm-thick undoped buffer layer, a Be-doped region, and a 1000-Å-thick undoped GaAs cap. The Be impurities are evaporated onto the GaAs epitaxial surface during an interruption of the epitaxial GaAs growth. After the Be-doping process, which lasts typically several seconds to several minutes (depending on the Be doping density), the epitaxial growth of GaAs is resumed. This growth-interrupted doping process allows one to obtain δ-function-like doping distributions provided that no spatial redistribution of impurities occurs during subsequent epitaxial growth. The Be-doping distributions are analyzed by secondary-ion-mass spectrometry (SIMS). This technique offers high sensitivity and, using low primary ion energies, high spatial resolution.

Be-doping profiles of four GaAs epitaxial layers are shown in Fig. 1. The two-dimensional doping densities range from 2×10^{13} cm^{-2} to 4×10^{14} cm^{-2}. At low densities the doping profile is spatially confined as expected for growth-interrupted or δ-doped semiconductors. The striking result of the experiment is obtained at the extremely high densities. As the density increases, the *peak concentration* reaches values of approximately 10^{20} cm^{-3}. This peak concentration *does not increase* even for higher impurity doses. Instead, the width of the doping profile increases (by more than a factor of 10) and is directly proportional to the two-dimensional (2D) doping density. The doping profile strongly resembles a top-hat distribu-

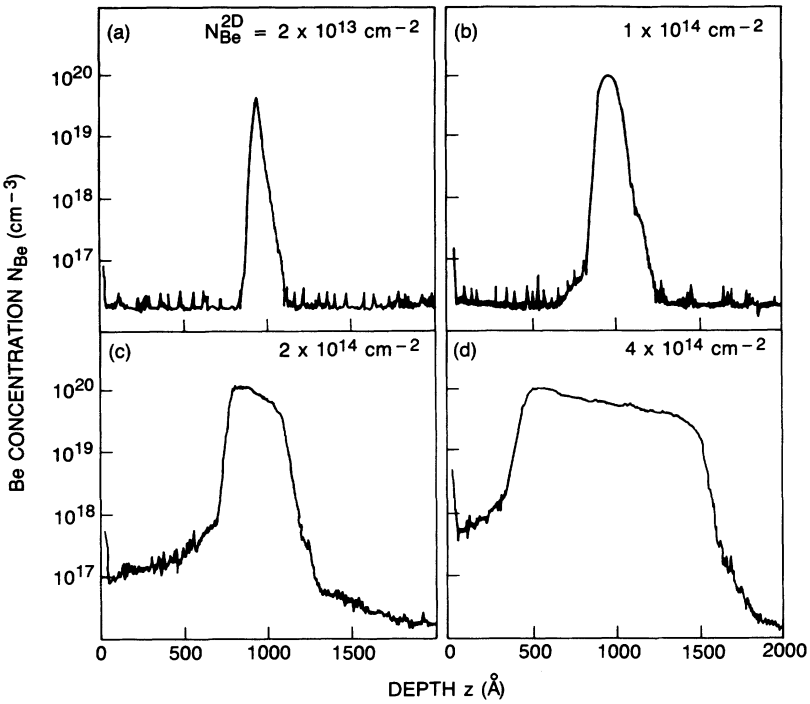

FIG. 1. Impurity profiles of GaAs-Be doped during growth interruption. The SIMS profiles exhibit a top-hat distribution that broadens with increasing two-dimensional Be density. The maximum concentration is approximately 10^{20} cm^{-3}.

tion whose width is given by $W = N^{2D}/N^{max}$, where N^{2D} and N^{max} are the 2D doping density and N^{max} is the peak concentration. Nearly all Be impurities were previously shown to be electrically active acceptors; the concentration of inactive Be impurities is negligibly small.[5] Figure 1(d) shows that impurities also diffuse into the already-grown material. However, this diffusion strongly depends on the impurity deposition time.

We next show that repulsive interactions between impurities can indeed limit the maximum concentration of impurities. The repulsive interaction between impurities is a result of an increase in total incorporation energy, as will be discussed later. We use the Monte Carlo method to model the impurity incorporation and diffusion processes during crystal growth. Impurities can occupy positions on the surface of the growing three-dimensional crystal and in the bulk. At the start of the simulation, the impurities are distributed at random positions on the surface where they are of neutral charge. The impurity coverage is uniform except for random local fluctuations in density. This corresponds to the experimental situation just after the deposition of the impurities. Then the surface is advanced at a rate corresponding to the crystal growth rate and, at the same time, the impurities are redistributed by Monte Carlo diffusion events.

In our calculation, impurity atoms are selected on a random basis and a possible diffusion hop is arranged. That is, a different position for the atom is chosen using a triplet of random numbers to obtain the change in the three cartesian coordinates. The second position is close to the original one, with the average displacement approximately equal to the distance between neighboring sites in the host lattice. The change in potential energy ΔE for the diffusion hop is calculated from Coulombic impurity interactions, which will be discussed below. Diffusion hops from the surface to the bulk and vice versa involve a change in the charge state, with no Coulomb effects for an atom on the surface. In general there can also be a term corresponding to differences in the local bonding interactions on the surface and in the bulk, but this term is set to zero in all of the simulations reported here (which applies to systems without conventional surface segregation). If the diffusion hop arranged for an atom in the bulk involves a changed position outside the crystal, it is projected back onto the crystal surface.

Not all of the diffusion events so chosen are actually executed by moving the atoms to the new positions. Successful events are chosen according to the Metropolis algorithm,[6] i.e., if ΔE is negative, reducing the total energy of the system, then the atom is moved, whereas if ΔE is positive, the event is executed with a probability $\exp(-\Delta E/kT)$. This insures that the system will ultimately approach an equilibrium configuration in the limit where the surface velocity approaches zero. Atoms on the surface are "trapped" or converted to bulk atoms. The flux of atoms incorporated by crystal growth is assumed to be $v_g N^s$, where v_g is the growth rate and N^s is the surface density. At temperatures ≥ 400 °C, the impurity atoms are likely to diffuse back to the surface, since the potential energy is lower for the uncharged surface impurity and the impurity diffusion coefficient is sufficiently large to make such a jump probable.

The calculated doping distributions with and without Coulombic interaction are shown in Fig. 2. The calculation uses the diffusion coefficient of Be at 500 °C.[5] The two profiles shown in Fig. 2 reveal a striking difference

between the profiles with and without repulsive interactions between impurities. The conventional profile has an approximately 10-times-higher peak concentration. In addition, the profile is approximately 10 times narrower. The profile obtained with repulsive effects has a rigid limit for the peak concentration and the shape of the profile resembles a top-hat distribution. Furthermore, the profile clearly resembles the experimental result for Be-doped GaAs. The incorporation of Be occurs at a concentration of 10^{20} cm^{-3} until the reservoir of Be atoms floating on the surface is exhausted.

The effect of different impurity densities is shown in Fig. 3. Inspection of the doping profiles shown in Fig. 3 reveals that the spreading is proportional to the impurity density. The peak impurity concentration increases slightly with the impurity dose. However, the peak concentration increases very sublinearly (about 20% for a 10-times-higher dose) with the impurity density.

It is important to note that the experimental and calculated results cannot be explained by a concentration-dependent diffusion coefficient or by surface segregation. A power-law dependence of the diffusion coefficient is predicted by the well-known substitutional-interstitial model.[7] However, calculated impurity profiles of this diffusion mechanism[7] did not display the top-hat distribution observed experimentally (see Fig. 1). Furthermore, surface segregation of impurities cannot explain the experimental results shown in Fig. 1. Exponentially decaying doping profiles rather than top-hat distributions result from impurity surface segregation.

The peak concentration obtained in the calculation is not universally valid for all impurities in semiconductors. Instead, the peak concentration is related to the diffusion coefficient. The redistribution of impurities with a very small diffusion coefficient (e.g., C in GaAs or B in Si) is negligible, and the observed distribution accurately reflects the incorporation of the impurities during crystal growth. On the other hand, a strongly diffusing impurity such as Zn in GaAs will redistribute more easily and be subjected to interaction effects at lower concentrations. The relation between the maximum experimentally achieved concentration and the diffusion coefficient is shown for acceptors in GaAs in Fig. 4. Figure 4 shows that the maximum experimentally achieved concentration decreases as the impurity-diffusion coefficient increases. This may be a result of enhanced segregation of the mobile impurities to the surface as well as redistribution in the bulk crystal.

We next discuss the physical basis of the repulsive impurity interaction. Consider a semiconductor whose Fermi level is pinned at the surface due to Bardeen states.[8] Such pinning occurs close to the middle of the fundamental gap in GaAs and persists even at growth temperatures such as 500–600 °C.[9] The pinning of the Fermi level re-

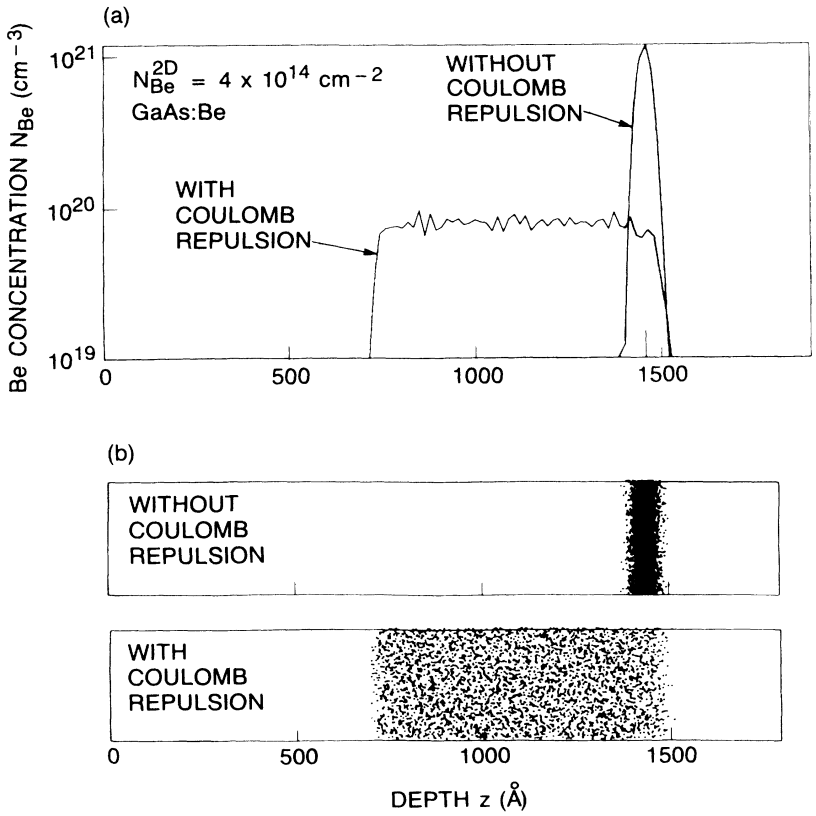

FIG. 2. (a) Calculated impurity distribution in a semiconductor doped during growth interruption with and without taking into account Coulomb correlation effects. A growth rate of 1 Å/sec and a diffusion coefficient of 2×10^{16} cm^2/sec is used in the calculation. (b) The actual distribution of impurity atoms is characterized by a maximum concentration.

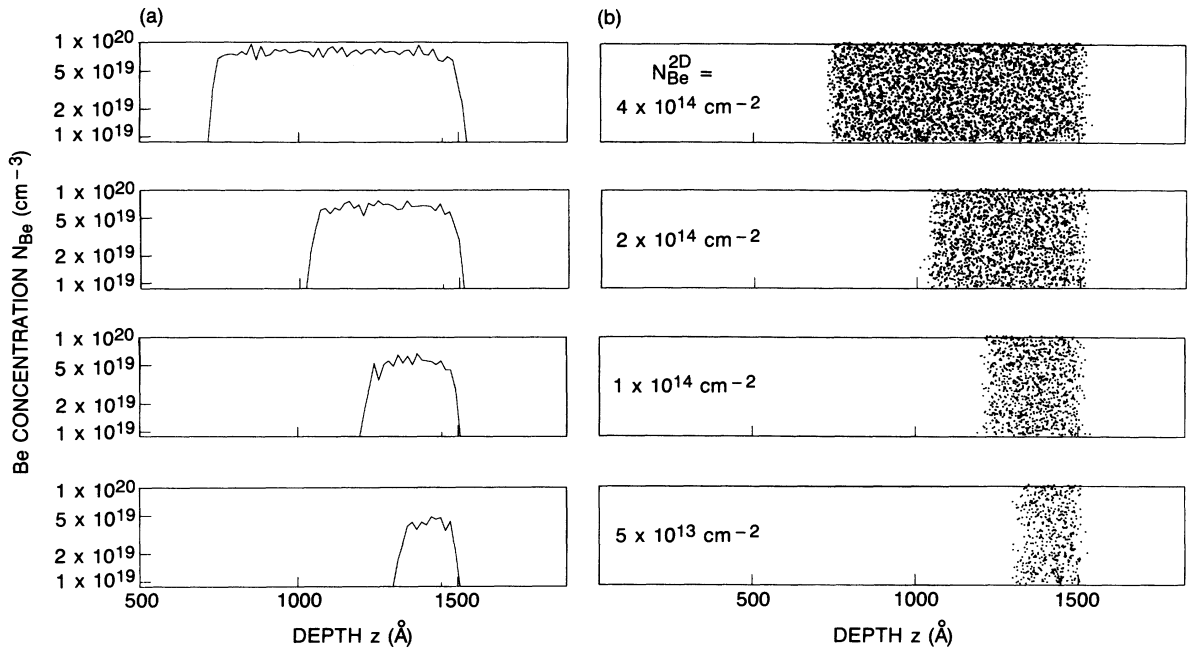

FIG. 3. (a) Calculated impurity distributions in a semiconductor doped during growth interruption with different doses of impurities. (b) Actual distribution of the calculated impurity atoms.

sults in a near-surface region that is depleted of free carriers. The resulting force due to Coulombic interactions is directed towards the surface, i.e., results in a movement of impurities toward the surface. The electric field at the semiconductor surface due to the charged impurities in the depletion layer can be obtained from Poisson's equation

$$\mathcal{E}=\sqrt{2eN\phi_B/\epsilon}\ , \tag{1a}$$

where N is the impurity concentration and ϕ_B is the barrier height of the pinned surface. Charged impurities experience a driving force towards the surface. The drift velocity of the impurities is given by $v_d=\mu\mathcal{E}$. Using the Einstein relation $\mu=eD/kT$, one obtains

$$v_d=\frac{eD}{kT}\sqrt{2eN\phi_B/\epsilon} \tag{1b}$$

where D and ϵ are the impurity diffusion coefficient and the permittivity, respectively. Thus, the Coulombic repulsion between impurities results in a redistribution of impurities that tends to reduce the impurity concentration. The maximum impurity concentration is obtained when the epitaxial growth rate v_g equals the drift velocity of impurities towards the surface, i.e., $v_d=v_g$. Using Eq. (1b), this condition yields

$$N^{\max}=\left[\frac{v_g kT}{eD}\right]^2\frac{\epsilon}{2e\phi_B}\ . \tag{2}$$

The concentration given by Eq. (2) is the maximum doping concentration achievable under the assumed conditions. The values for $N^{\max}$ calculated from Eq. (2) are in amazingly good agreement with experimental values. As an example, we consider Be-doped GaAs with $D\cong 3\times 10^{-16}$ cm^2/sec, $v_g=1$ $\mu m/h$, $\phi_B=0.9$ V, $T=800$ K, and calculate from Eq. (2) $N_{Be}^{\max}\cong 1.6\times 10^{20}$ cm^{-3} in good agreement with the experimental value. Furthermore, Eq. (2) predicts a v_g^2 dependence of the maximum impurity concentration on the growth rate. However, such a dependence was not yet confirmed in our experiments due to the well-known difficulty of changing the growth rate over a wide range.

An increase in the total incorporation energy also results from phase-space filling, i.e., an increase in electron-

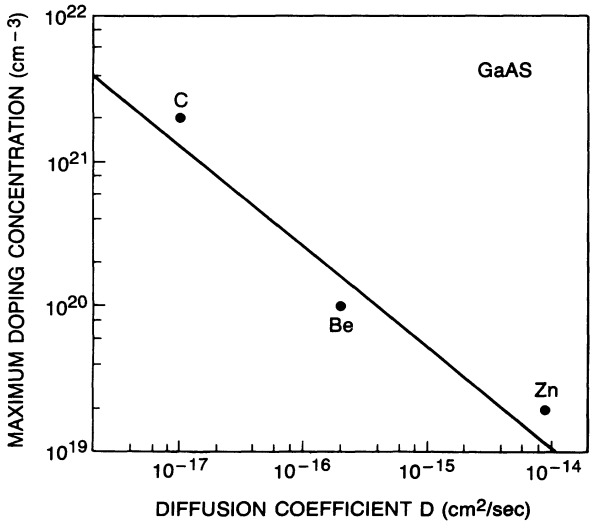

FIG. 4. Maximum experimentally achieved impurity concentration as a function of the impurity diffusion coefficient for different acceptors (C, Be, Zn) in GaAs. The solid line is a least-square fit to the experimental points.

ic energy. The increase in Fermi energy with impurity concentration in a degenerately doped, single-spherical-valley semiconductor is given by

$$E_F - E_C = \left[\frac{3}{4}\sqrt{\pi}\frac{N}{N_C} \right]^{2/3}, \tag{3}$$

where N_C is the effective density of states at the band edge. The electronic energy increases with doping density and can assume values of several hundred meV at high n-type impurity concentrations in GaAs. The repulsive interaction between impurities can be expressed also in terms of Yukawa potentials

$$E = \frac{1}{2}\sum_{i \neq j} \frac{e^2}{4\pi\epsilon r_{ij}} e^{-r_{ij}/r_s} \tag{4}$$

where r_s is the screening radius which, in the degenerate limit, does not depend strongly on the free-carrier concentration. Assuming that the impurities are randomly distributed, it can be shown that Eqs. (3) and (4) have the *identical* functional dependence on the impurity concentration[10] and both equations yield energies of similar magnitude.[10] Therefore Eqs. (3) and (4) may describe the same physical phenomenon by different approaches. However, more work is needed to prove the equivalence of the two approaches.

Finally, we compare the strength of the Coulomb repulsions with other factors that may limit impurity concentrations using molecular-dynamics calculations. A crucial aspect of this comparison is the dependence of the impurity chemical potential μ_c on concentration N. A steep increase in μ_c with N would favor a concentration limit that is relatively independent of growth conditions. First, we note that the equilibrium solubility limit $N^{(0)}$ is the concentration at which the impurity chemical potential is equal to that of impurity precipitates μ_p; i.e., $\mu_c(N^{(0)}) = \mu_p$. In the case of noninteracting impurities, the system behaves as an ideal gas and μ_c increases as $kT\ln(N)$. The chemical potential increases by only $2.3kT$ for an order-of-magnitude increase in N. Impurity concentrations much greater than $N^{(0)}$ cause only a small increase in μ_c over μ_p.

Most impurities will produce a local strain field, and these interactions should be considered. Substitutional Be in GaAs is expected to create a region of tensile stress because of the smaller atomic radius of Be. These stress fields are the source of repulsive interactions similar to Coulomb fields. We have examined an atomic-scale model of silicon in order to assess the strength and range of these interactions. The Stillinger-Weber interatomic potential was used to approximate the interactions between silicon atoms.[11] Elastic constants for the diamond-cubic crystal composed of these atoms have been measured, and are within 30–40 % of the experimental measurements for silicon.[12] The impurity atoms were assigned interactions of the Stillinger-Weber form. But the atomic-size parameter σ_{ih} for the impurity-host interaction is scaled up or down relative to the host-host parameter σ_{hh} to account for larger or smaller atomic diameters.

The model system was prepared in the shape of a thin film. Periodic boundary conditions were applied on the four sides of a rectangular computational cell that were perpendicular to the film surfaces; this accounts for an infinite lateral extent of the film and prevents expansion or contraction of the film lattice constant in these directions. Expansion in the direction perpendicular to the film surfaces is, of course, permitted because these surfaces are not constrained. The films consisted of 31 layers parallel to the (100) surfaces, and the impurities occupy sites in the central (100) layer. All atoms in the initial configuration are located at exact diamond-cubic lattice sites. Strain fields around the impurities developed as the system was relaxed to an equilibrium state using molecular-dynamics methods. Briefly, the particles are allowed to move according to classical equations of motion, but in the presence of dissipative forces. The unbalanced forces on particles in the initial configuration cause accelerations toward positions of lower potential energy and the dissipation absorbs the kinetic energy and allows the system to relax toward the local minimum in potential energy.

The impurity atom energies E_i calculated by this method are given in Table I. This energy is defined as the difference between the total energy of the system containing the impurity atoms and that of a reference system in which the impurities are replaced by host atoms. The fraction θ of the (100) layer occupied by impurities is given in the first column. In every case the impurities occupy a square lattice; with $\theta = 1$, every (100) site is occupied by an impurity, $\theta = 0.25$ has impurities at twice the interatomic spacing of the (100) atoms, etc. Over the range of impurity sizes listed in the table, the impurity energy scales as the square of the misfit, or $(1 - \sigma_{in}/\sigma_{hh})^2$. This indicates that the atomic positions remain roughly within the harmonic regions of the interatomic potential, even with 10% atomic size differences.

The most striking aspect of these results is the small effect of impurity-impurity interactions on their energies. If we ignore the data for $\theta = 1$, the impurity-impurity interaction energy scales with inverse distance r_{ii}; i.e., $E_i = E_0 + E_1/r_{ii}$, where E_0 and E_1 are constants for a

TABLE I. Impurity-atom energies E_i.

σ_{ih}/σ_{hh}	θ	E_i
0.99	1.0	0.003 58
0.99	0.25	0.003 27
0.96	1.0	0.058 4
0.96	0.25	0.055 5
0.96	0.062 5	0.051 0
0.90	1.0	0.315
0.90	0.25	0.349
0.90	0.0625	0.325
0.90	0.015 625	0.311 5035
1.10	1.0	0.269 8
1.10	0.25	0.305 5
1.10	0.062 5	0.285 9
Coulomb	1.0	5.352
Coulomb	0.25	1.213
Coulomb	0.062 5	0.248 4

given impurity type. An extrapolation of the results for $\sigma_{ih}/\sigma_{hh}=0.9$ to $r_{ii}=\infty$ yields a value $E_0=0.302$ eV, and this accounts for the major part of E_i even when $\theta=1$. The r_{ii} dependence indicates that the strain field of an impurity falls off as $1/r_{ii}$, and that the field extends only to the nearest-neighbor shell of impurities. The screening at the nearest-neighbor position is a result of relaxation at the free surface and lateral constraints imposed by coherent growth on a thick substrate in the experiments. (The boundary conditions at the lateral computational cell walls have the equivalent effect in the simulations.) The anomalous $\theta=1$ energies are caused by atomic-scale effects that are specific to the (100) structure. Atoms in (100) layers have no nearest neighbors within the layer. Thus, all impurity-host bonds extend between the layer with the impurities and the layer above or the layer below. When all atoms in the central (100) layer contain impurities, these bonds can therefore lengthen or contract to the optimum length, since the free surfaces permit the crystal regions bounding the impurity layer to move without constraints. Such relaxations will affect the bond angles, but the absence of the bond-stretching forces causes the energy of this configuration to be lower than expected.

The situation is quite different in the case of Coulomb forces. In this case the screening distance is relatively insensitive to N, and for comparison we have also listed in Table I the screened Coulomb energies for impurities at the same positions as those used in the molecular-dynamics calculations. Here we used $\lambda_s=25$ Å and $\epsilon=13$ for all impurity densities (since the impurities are confined to a plane, the density dependence should be even slower than $N^{1/6}$). In this case the impurity energies are larger than that for impurities with $\sigma_{ih}/\sigma_{hh}=0.9$, and the density dependence is much stronger. The values of E_i scale roughly as $1/r_\mu^2$, corresponding to the amount of charge within the screening distance of a given impurity. It is clear from these results that Coulomb interactions cause a much steeper increase in μ_i with N than do the strain fields of the impurities.

The consideration of repulsive interaction energies gives rise to a number of intriguing questions. It was, for example, previously assumed that impurities are distributed randomly in semiconductors.[13] The Coulomb correlation effects by definition result in a more ordered non-Poissonian distribution. Ordered impurity distributions are expected to result in significant modifications of the transport characteristics.[14] The minimization of the Coulombic interaction is obtained if impurities are arranged in the face-centered-cubic lattice. The ordered arrangement of impurities in such a lattice requires a novel understanding of the transport properties in extremely highly doped semiconductors. The ordered, i.e., periodic superlattice potential does not scatter the charge carriers propagating in the potential. Thus, the propagation of carriers without impurity scattering would be possible in highly doped semiconductors.

Furthermore, the random distribution of impurities results in many well-known phenomena such as luminescence line broadening, electronic screening, hopping conduction, etc. The ordered distribution of impurities substantially modifies such phenomena. Our understanding of impurities in semiconductors needs substantial modification in the extreme doping regime.

In conclusion, repulsive correlation between impurities are investigated at extremely high doping concentrations in semiconductors. The repulsive Coulomb interaction in the surface depletion region of a semiconductor results in impurity drift along with the growth surface. It is shown that this mechanism can limit the attainable maximum impurity concentration in semiconductors. Monte Carlo calculations are employed to model the impurity interaction. The calculations confirm the limitation of the maximum impurity concentration due to repulsive interactions. Experimental results on Be-doped GaAs indeed exhibit a rigid limit of the maximum Be-impurity concentration of approximately 10^{20} cm^{-3}. This experimental value is in agreement with model calculations. Finally, we show that the maximum impurity concentration is expected to increase with a decreasing diffusion coefficient. This trend is compared with the maximum concentration of C, Be, and Zn in GaAs. Even though a clear functional dependence of the maximum impurity concentration on the impurity diffusion coefficient cannot be established, the qualitative trend of experiment and theory are in agreement.

---

[1]The solid solubility limit for impurities in semiconductors are quite well documented for silicon. See, for example, *Properties of Silicon* (INSPEC, New York, 1988) or *Quick Reference Manual,* edited by W. E. Beadle, J. C. C. Tsai, and R. D. Plummer (Wiley, New York, 1985).

[2]The impurity concentration achievable in δ-doped semiconductors and during low-temperature growth allow one to significantly exceed the solid-solubility limit. See, for example, E. F. Schubert, B. Ullrich, T. D. Harris, and J. E. Cunningham, Phys. Rev. B **38**, 830 (1988); H. J. Gossmann, E. F. Schubert, D. J. Eaglesham, and M. Cerullo, Appl. Phys. Lett. **57**, 2440 (1990).

[3]T. N. Theis, P. M. Mooney, and S. L. Wright, Phys. Rev. Lett. **60**, 361 (1988).

[4]G. A. Baraff and M. Schlüter, Phys. Rev. Lett. **55**, 1327 (1985); W. Walukiewicz, Appl. Phys. Lett. **54**, 2094 (1989); E. Tokumitsu, J. Appl. Phys. (to be published).

[5]E. F. Schubert, J. M. Kuo, R. F. Kopf, H. S. Luftman, L. C. Hopkins, and N. J. Sauer, J. Appl. Phys. **67**, 1969 (1990).

[6]R. W. Hockney and J. W. Eastood, in *Computer Simulation using Particles* (McGraw-Hill, New York, 1981).

[7]F. C. Frank and D. Turnbull, Phys. Rev. **104**, 617 (1956). L. R. Weisberg and J. Blanc, Phys. Rev. **131**, 1548 (1963).

[8]J. Bardeen, Phys. Rev. **71**, 717 (1947).

[9]Y. Nanichi and G. L. Pearson, Solid-State Electron. **12**, 341 (1969).

[10]G. H. Gilmer and E. F. Schubert (unpublished).
[11]F. H. Stillinger and T. A. Weber, Phys. Rev. B **31**, 5262 (1985).
[12]M. D. Kluge, J. R. Ray, and A. Rahman, J. Chem. Phys. **85**, 4028 (1986).
[13]W. Shockley, Solid-State Electron. **2**, 35 (1961).
[14]R. L. Headrick, I. K. Robinson, E. Vlieg, and L. C. Feldman, Phys. Rev. Lett. **63**, 1253 (1989); see also A. F. J. Levi, S. L. McCall, and P. M. Platzman, Appl. Phys. Lett. **54**, 940 (1989).