CHAPTER 9: Ion Implantation

Ion implantation is a low-temperature technique for the introduction of impurities (dopants) into semiconductors and offers more flexibility than diffusion. For instance, in MOS transistors, ion implantation can be used to accurately adjust the threshold voltage.

In ion implantation, dopant atoms are volatilized, ionized, accelerated, separated by the mass-to-charge ratios, and directed at a target that is typically a silicon substrate. The atoms enter the crystal lattice, collide with the host atoms, lose energy, and finally come to rest at some depth within the solid. The average penetration depth is determined by the dopant, substrate materials, and acceleration energy. Ion implantation energies range from several hundred to several million electron volts, resulting in ion distributions with average depths from < 10 nm to 10 μ. Doses range from 10^{11} atoms/cm^2 for threshold adjustment to 10^{18} atoms/cm^2 for buried dielectric formation.

9.1 Ion Stopping

As each implanted ion impinges onto the target, it undergoes a series of collisions with the host atoms until it finally stops at some depth, as depicted in ***Figure 9.1***. Since the initial ion energy, typically several tens of keV, is much higher than lattice binding energies, the ion scattering process can be simulated based on elastic collisions between pairs of nuclei while ignoring the relatively weak lattice forces. A second component of scattering comes from inelastic collisions with electrons in the target. The total stopping power *S* of the target, defined by the energy loss (*E*) per unit path length (*x*) of the ion, is the sum of these two terms:

$$S = \left(\frac{dE}{dx}\right)_{nuclear} + \left(\frac{dE}{dx}\right)_{electronic} \qquad \textit{(Equation 9.1)}$$

Figure 9.2 shows the relative distribution to *S* of each of the terms over a wide energy range. Energies typical for ion implantation, 10 to 200 keV, fall at the far left of the figure, a region dominated by nuclear stopping.

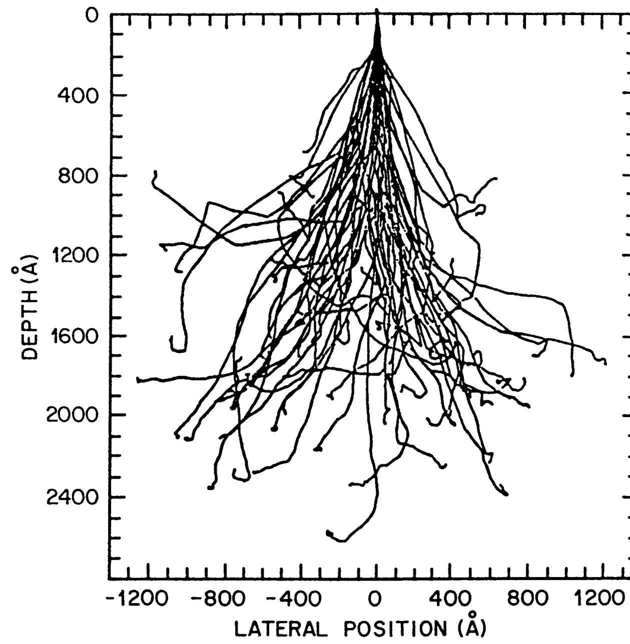


Figure 9.1: Monte Carlo calculation of 128 ion trajectories for 50 keV boron implanted into silicon.

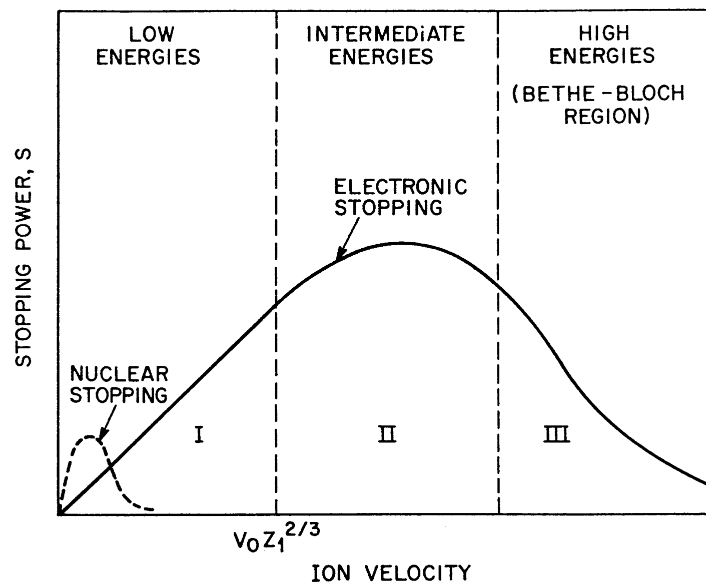


Figure 9.2: Nuclear and electronic components of the ion stopping power as a function of ion velocity. The quantity v_o is the Bohr velocity, $\frac{q^2}{4\pi\varepsilon_o\hbar}$, and Z_1 is the ion atomic number.

Nuclear stopping is caused by a collision between two atoms, and can be described by classical kinematics. If the atoms were bare nuclei, then at a separation *r*, the coulombic potential between them would be:

$$V_c(r) = \frac{q^2 z_1 z_2}{4\pi\varepsilon_o r} \qquad \textit{(Equation 9.2)}$$

where z_1 and z_2 are the atomic number of the implanted and target atoms, respectively, ε_o is the permittivity, and q is the electronic charge. In reality, electrons screen the nuclear charge and a screening function, $f_s(r)$, must be included such that:

$$V(r) = V_c(r)\, f_s(r) \qquad \textit{(Equation 9.3)}$$

Given the interaction potential, the equations of motion of atoms can be integrated to yield the scattering angle for any incident ion trajectory. Working in the center-of-mass frame, the result is:

$$T = \frac{4M_1M_2}{(M_1+M_2)^2} E \sin^2\left\{\frac{\theta}{2}\right\} \qquad \textit{(Equation 9.4)}$$

where T is the energy lost by the incoming ion, E is the energy of the ion, θ is the scattering angle, and M_1 and M_2 are the atomic mass numbers of the ion and target atom, respectively. The rate of energy loss to nuclear collisions per unit path length can be calculated by summing the energy loss multiplied by the probability of that collision occurring. If the maximum possible energy transfer in a collision is T_{max} and there are N target atoms per unit volume, then

$$S_n = \left(\frac{dE}{dx}\right)_{nuclear} = N\int_0^{T_{\max}} T d\sigma \qquad \textit{(Equation 9.4)}$$

where $d\sigma$ is the differential cross section. Nuclear stopping is elastic, and so energy lost by the incoming ion is transferred to the target atom that is subsequently recoiled away from its lattice site, thus creating a damage or defect site.

Electronic stopping is caused by the interaction between the incoming ion and the electrons in the target. The theoretical model is quite complex, but in the low energy regime, the stopping is similar to a viscous drag force and is proportional to the ion velocity. Electronic stopping is inelastic. The energy loss by incident ions is dissipated through the electron cloud into thermal vibrations of the target.

9.2 Range Distributions

Each implanted ion traverses a random path as it penetrates the target, losing energy by nuclear and electronic stopping. Since implantation doses are usually higher than 10^{12} ions/cm^2, ion trajectories can be predicted employing statistical means. The average total path length is called the range R, which is composed of both lateral and vertical motions. The average depth of the implanted ions is called the projected range R_p, and the distribution of the implanted ions about that depth can be approximated as Gaussian with a standard deviation σ_p (or ΔR_p). The lateral motion of the ions leads to a lateral Gaussian distribution with a standard deviation $\sigma_\perp$. These parameters are illustrated schematically in ***Figure 9.3***. Far from the mask edge, the lateral motion can be ignored, and $n(x)$, the ion concentration at depth x, can be written as:

$$n(x) = n_o \exp\left\{-\frac{(x-R_p)^2}{2\sigma_p^2}\right\} \qquad \textit{(Equation 9.6)}$$

where n_o is the peak concentration, R_p is the projected range, and σ_p is the standard deviation. If the total implanted dose is ϕ, integrating *Equation 9.6* gives an expression for the peak concentration n_o:

$$n_o = \frac{\phi}{\sqrt{2\pi}\sigma_p} \cong \frac{0.4\phi}{\sigma_p} \qquad \textit{(Equation 9.7)}$$

In general, an arbitrary distribution can be characterized in terms of its moments. The normalized first moment of an ion distribution is the projected range, R_p. The second moment is the standard deviation, σ_p. The third moment is the skewness, γ, whereas the fourth moment, kurtosis, is designated β. Qualitatively, skewness is a measure of the asymmetry of the distribution. Positive skewness places the peak of the distribution closer to the surface than R_p. Kurtosis is an indication of how flat the top of a distribution is. A true Gaussian distribution has a skewness of 0 and a kurtosis of 3.

Several different distributions have been employed to give a more accurate fit to the moments of an ion implant distribution than is possible using a Gaussian. The most popular of these is the Pearson IV fit. ***Figure 9.4*** compares experimental boron profiles under non-channeling conditions with fitted distributions for energies between 30 keV and 800 keV. As the energy is increased, the profiles become more negatively skewed and deviate more

significantly from a true Gaussian. For heavier ions (e.g. arsenic), the profiles at low energies have a positive skewness which decreases more slowly but can also become negative for sufficiently high energies. The nature of this skewness can be explained by increased electronic stopping for faster moving ions, that is, in the "pre - peak" or "pre - R_p" region. ***Table 9.1*** lists the range and standard deviation data for boron in a variety of materials.

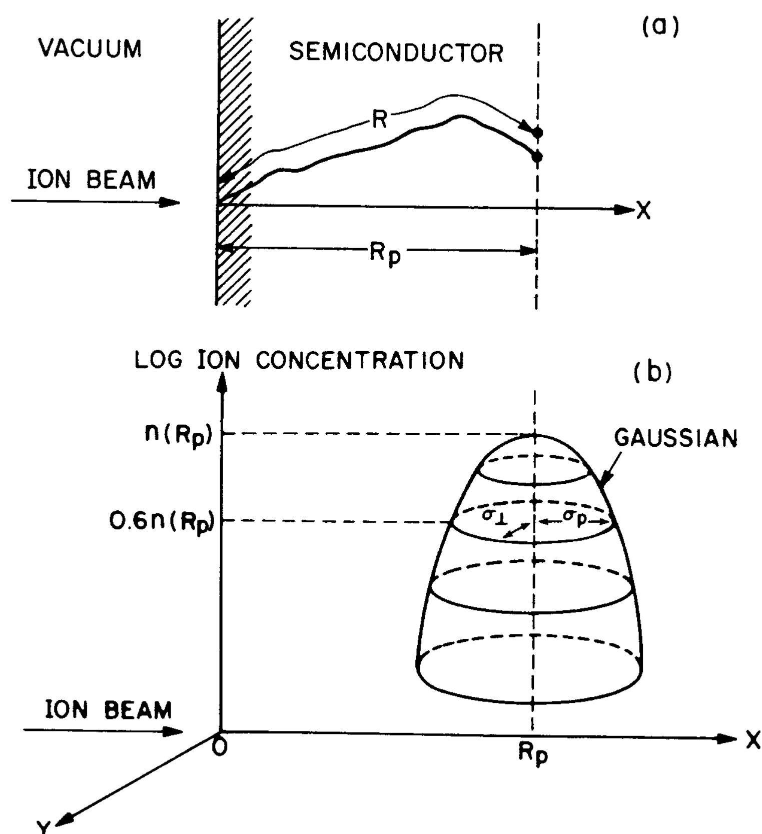


Figure 9.3: Schematic views of the ion range. (**a**) The total path length R is longer than the projected R_p. (**b**) The stopped atom distribution is two-dimensional Gaussian.

CityU
香港城市大學
City University of Hong Kong

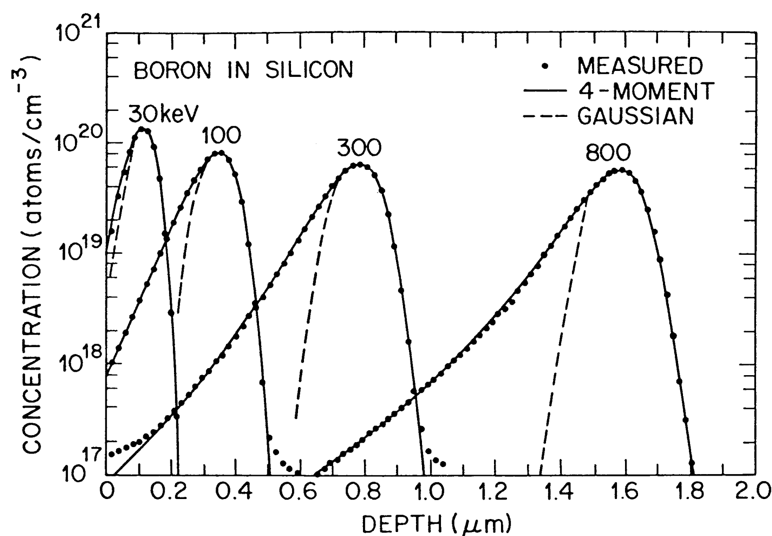


Figure 9.4: Boron implanted atom distributions, comparing measured data points with four-moment (Pearson IV) and Gaussian fitted distributions. The boron was implanted into amorphous silicon without annealing.

Table 9.1: Boron ranges in various materials.

100 keV boron implantation				
Material	Symbol	Density (g/cm^3)	R_p(Å)	σ_p(Å)
Silicon	Si	2.33	2968	735
Silicon dioxide	SiO_2	2.23	3068	666
Silicon nitride	Si_3N_4	3.45	1883	408
Photoresist AZ111	$C_8H_{12}O$	1.37	10569	1202
Titanium	Ti	4.52	2546	951
Titanium silicide	$TiSi_2$	4.04	2154	563
Tungsten	W	19.3	824	618
Tungsten silicide	WSi_2	9.86	1440	555

Experimentally, it is relatively easy to unveil vertical atomic profiles. However, it is much more difficult to accurately assess lateral atomic profiles, for example, at the edge of a mask, and so simple Gaussian distributions are frequently employed as an approximation. A two-dimensional distribution can be written as the product of vertical and lateral distributions:

$$n(x,y) = \frac{n_{vert}(x)}{\sqrt{2\pi}\sigma_{\perp}} \exp\left(\frac{-y^2}{2\sigma_{\perp}^2}\right) \qquad \textit{(Equation 9.8)}$$

This equation describes the result of implanting at a single point on the surface. To obtain the result of implanting through a mask window, *Equation 9.8* can be integrated over the open areas where the ion beam can enter. ***Figure 9.5*** displays the results for a 70 keV boron implant through a 1 μm slit in a thick mask showing that ions scatter well outside the open area. To minimize lateral scattering, masking layers are often tapered at the edge rather than perfectly abrupt, so that ions are gradually prevented from entering the silicon.

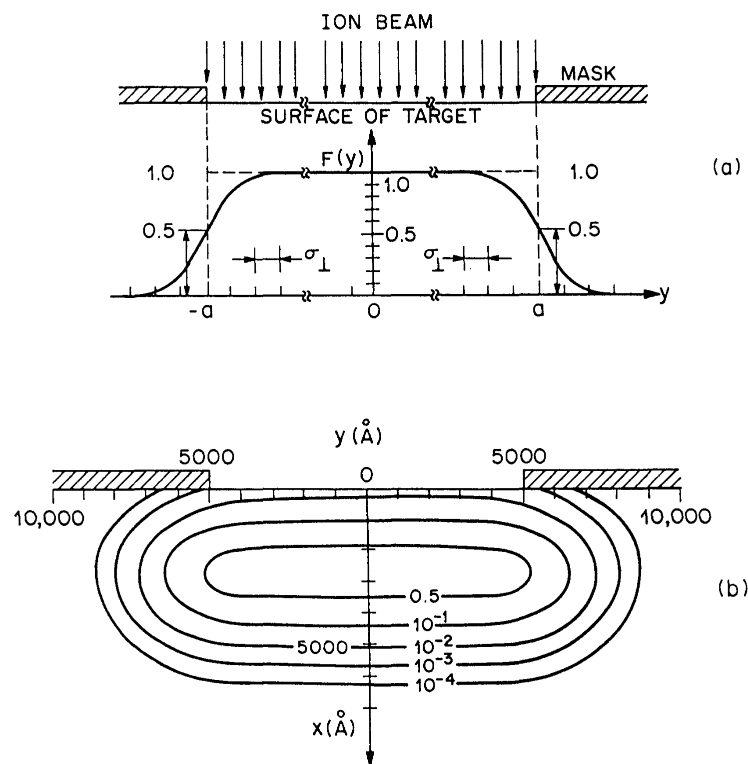


Figure 9.5: Two-dimensional implantation profiles. (**a**) Fraction of total dose as a function of lateral position for an opaque mask. (**b**) Equi-concentration contours for a 70 keV boron implant through a 1 μm slit.

9.3 Damage

As each ion penetrates the target, it undergoes a series of collisions displacing host atoms along the way. Both the ion and dislodged target atoms can continue and cause further damage, and so the energy is spread over many moving particles. Eventually, the energy per particle becomes too small and the cascade stops. Hence, after many ions have been implanted, an initially crystalline target will be so perturbed that it will have changed to a highly disordered state. If the target temperature is sufficiently high, the competing process of self-annealing can occur to repair some or all of the damage as it is generated. A critical dose can be defined as the minimum necessary to amorphize the target. The temperature dependence of the critical amorphization dose is exhibited in ***Figure 9.6***. Heavy ions displace a greater volume of target atoms per ion, and so a higher temperature is necessary for complete recrystallization.

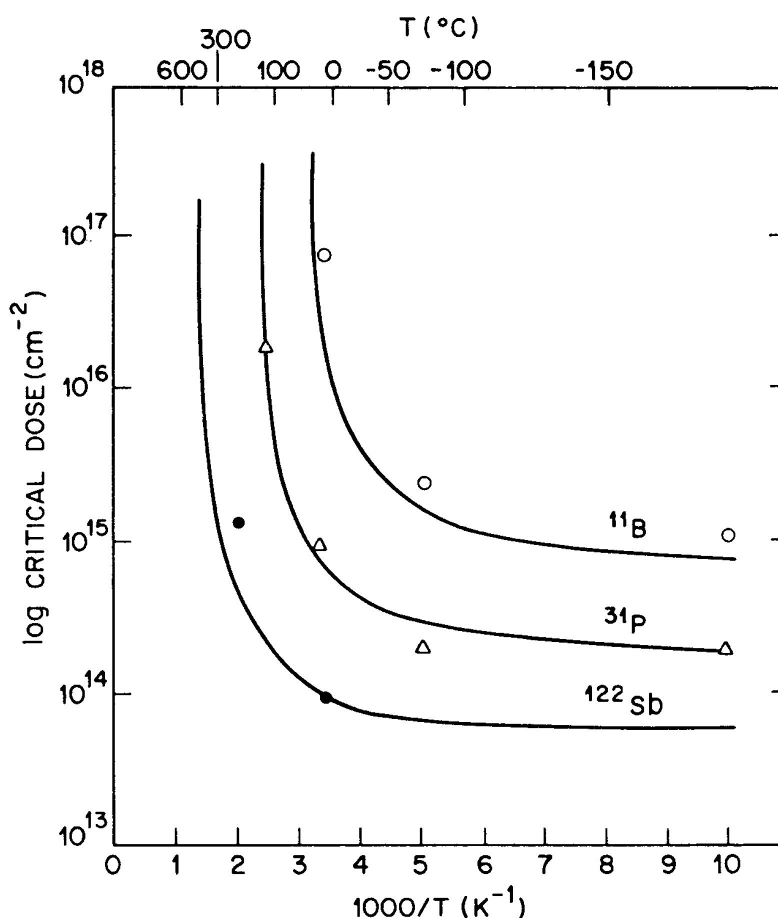


Figure 9.6: A plot of the critical dose necessary to make a continuous amorphous layer as a function of temperature.

Damage can affect the results of subsequent processing steps. For example, point defects are known to influence diffusion in silicon. Damaged oxide layers etch faster than undisrupted oxide because some of the bonds are already broken. Dangling oxide bonds can act as electron traps in finished devices, and the organic chains in photoresist films can be broken down into a carbon-rich material which is much harder to remove.

9.4 Channeling

For ions moving in certain directions in a crystalline material, there are long-range open spaces through which the ions can travel without significant scattering (***Figure 9.7***). Ions can be steered down these open channels by glancing collisions with the atom rows or planes, thereby extending the final ion distribution deeper into the target. ***Figure 9.8a*** shows three ions entering a simple cubic lattice. Ion A is well aligned with a channel and so suffers only glancing collisions with the walls as it travels far into the lattice. Ion B is scattered into a channel after a short distance. Ion C is not channeled and undergoes random collisions with lattice atoms. Experimentally, the angular width of a channel can be found by measuring the number of ions backscattered from the lattice as a function of angle for a well collimated beam (***Figure 9.8b***). The visible consequence of channeling is an added tail to the distribution (***Figure 9.8c***). The magnitude and extent of this tail depend on the mass and energy of the ions and the crystal orientation.

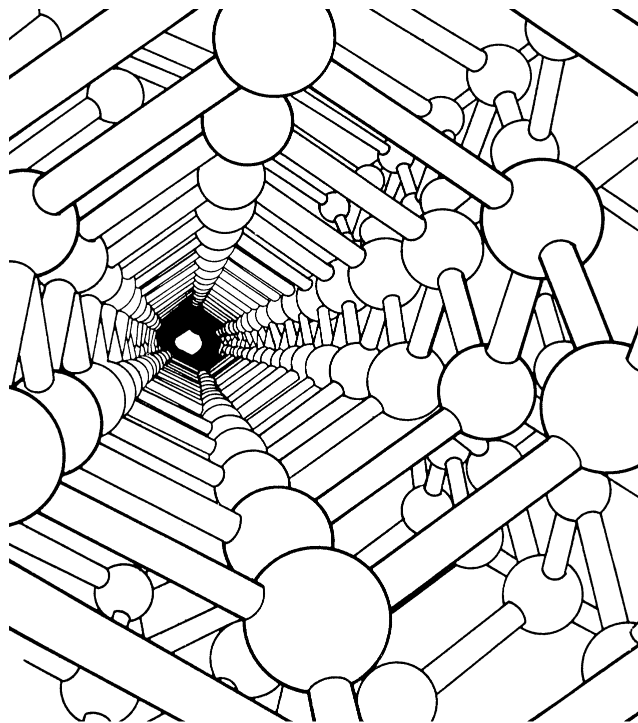


Figure 9.7: Model for a diamond structure, viewed along a <110> axis.

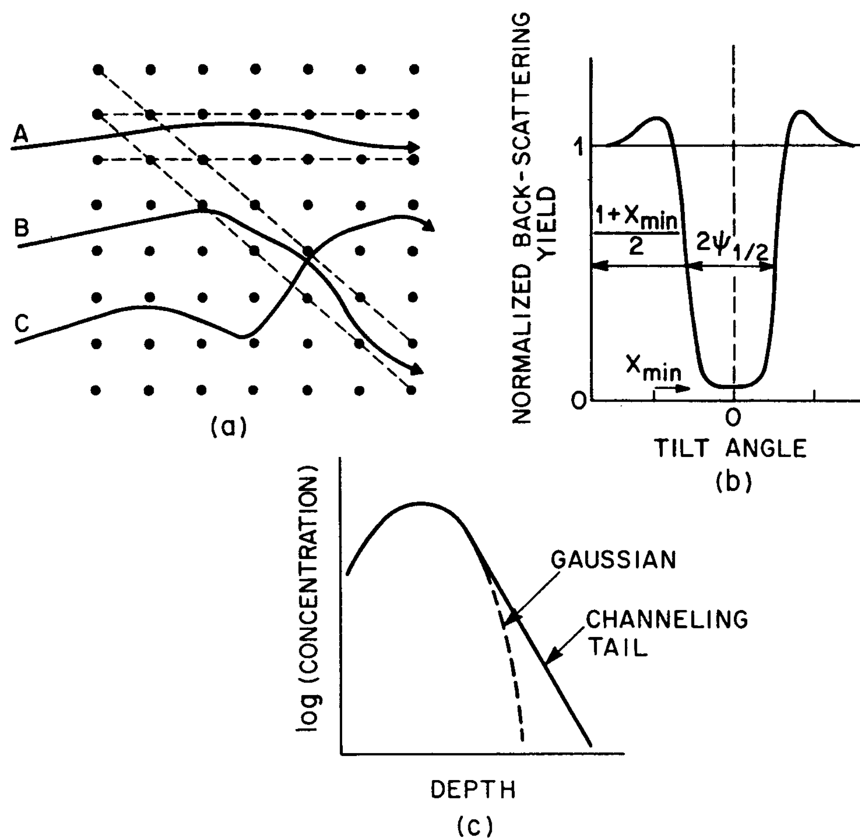


Figure 9.8: Schematic views of channeling. (**a**) Ion paths through a cubic lattice, showing channeled and nonchanneled cases. (**b**) Back-scattered yield around a channeling direction. Yield is a minimum when the beam is well aligned with a channel. (**c**) The effect of channeling is to add a tail to the atom distribution.

Channeling is characterized by a critical angle, ψ_1, which is the maximum angle between ion and channel for a glancing collision to occur. If we neglect thermal vibrations,

$$\psi_1 = 9.73\,(Z_1Z_2/Ed)^{1/2} \quad \text{(degrees)} \qquad \textit{(Equation 9.9)}$$

where E is the ion energy (keV) and d is the atomic spacing (A). *Equation 9.9* reveals that channeling is more likely for heavier ions (higher Z) and lower energy.

In order to preclude haphazard channeling, wafers are usually tilted by 7° to avoid the major planes. At first, it might seem that channeling provides an easy means for placing ions deep inside the target while minimizing lattice damage.

Unfortunately, it is not viable in practice. The channeled atom profile is sensitive to changes on the order of 1° in wafer tilt and beam divergence (***Figure 9.9***), and ions are scattered by amorphous surface films and residual damage from previous processing steps.

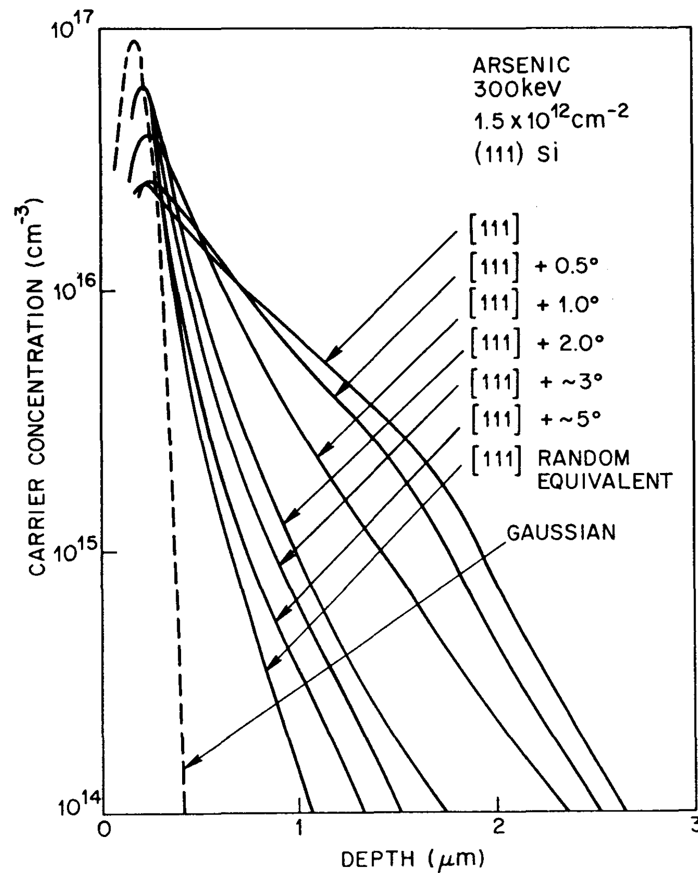


Figure 9.9: Electrically active arsenic distribution as a function of beam angle. The "random equivalent" case is the usual 7° tilt used to avoid channeling, and still shows significant differences from a Gaussian.

9.5 Recoils

If the target is composed of two layers, atoms will be displaced from one layer into the underlying layer during the implantation process. Depending on the materials involved, this phenomenon can be either useful or deleterious. The most common case, implantation through a thin oxide layer into silicon, is shown in ***Figure 9.10***. Oxygen atoms are dislodged from the oxide matrix and recoiled into silicon, giving a profile of two roughly exponential regions. Close to the interface, the displacement cascades push away very-low-energy oxygen atoms just across the interface. Beyond this, primary oxygen recoils give rise to a deeper exponential tail. This high interfacial oxygen concentration can degrade carrier mobility and introduce deep level traps.

Recoil mixing can be utilized positively to introduce dopant atoms that cannot be readily made into a source for implantation machines but can be deposited as a thin film on a silicon substrate. Implanting silicon through the film will push a dopant tail into silicon. Self-implantation can also be utilized to break up thin contaminant layers under deposited films prior to further processing, such as the formation of a metal silicide.

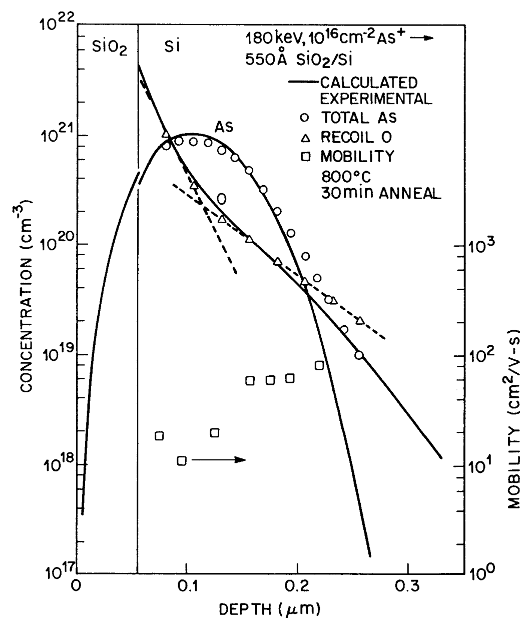


Figure 9.10: Oxygen recoils from the implantation of arsenic through an oxide layer. Solid lines are from Boltzmann calculations. Dashed lines indicate the two exponential regions of the recoil distribution.

9.6 Instrumentation

Figure 9.11 displays the schematic of a medium-energy ion implanter. The total number of ions entering the target is called the dose or fluence. If the current in the ion beam is I, then for a beam swept over an area A, the dose ϕ is given by:

$$\phi = \frac{1}{QA}\int I dt \qquad \text{(Equation 9.10)}$$

where Q is the charge on the ion and t is the implantation time. For an accurate assessment of the current, it is imperative that secondary electrons emitted from the target must be recaptured, for instance, by using a small positive bias around the Faraday cup cage.

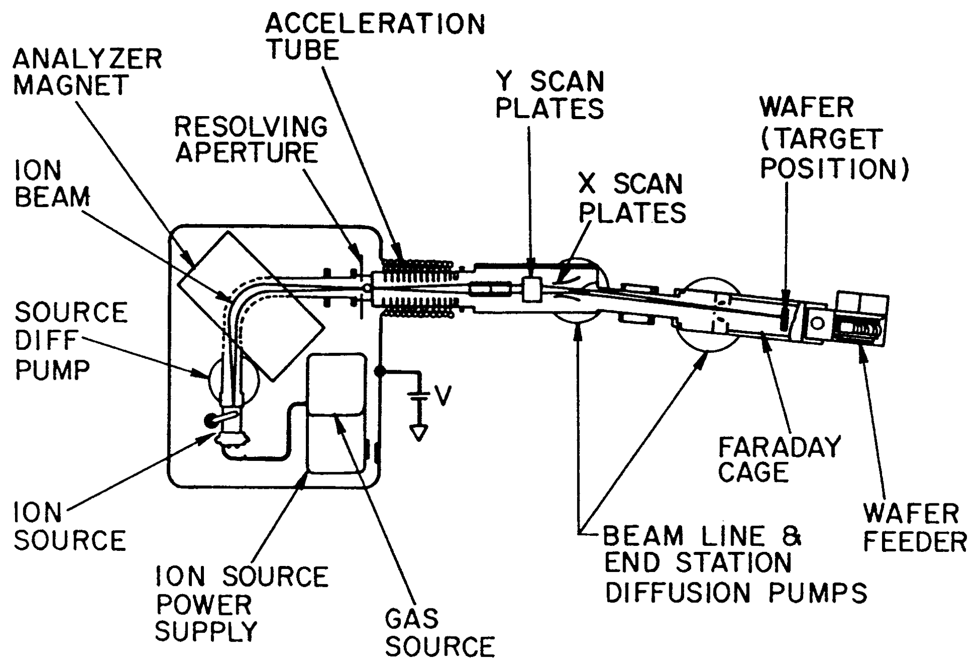
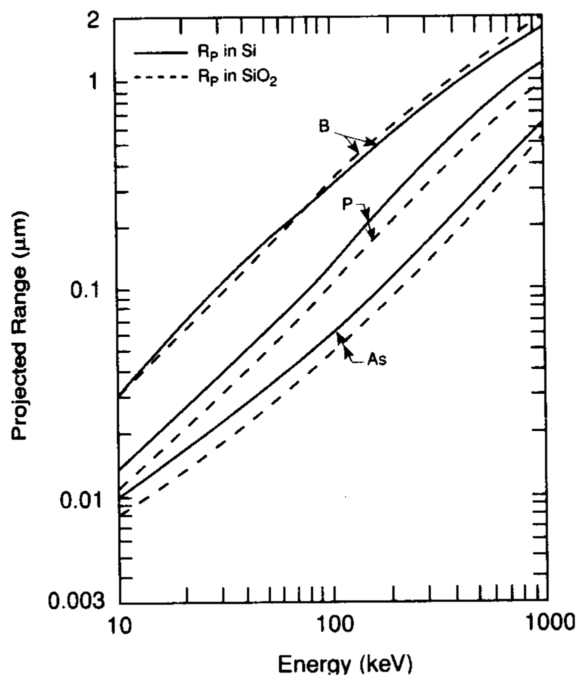


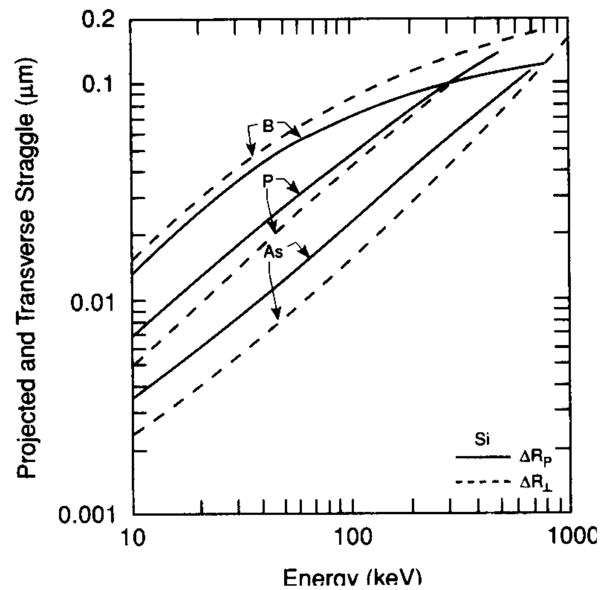
Figure 9.11: Schematic of a typical commercial ion implanter.

Example 9.1

A silicon wafer, 200mm in diameter, is implanted with 100 keV boron ions to a dose of 5 x 10^{15} cm^{-2}. Determine the projected range (R_p), projected straggle or standard deviation (σ_p), and peak concentration (n_o), assuming that the implant distribution is Gaussian. Calculate the required beam current if the implantation time is 90 seconds.



Projected range (R_p) for B, P, and As in Si and SiO_2 at various implantation energies



Ion projected straggle or standard deviation (ΔR_P or σ_p) for As, P, and B in silicon

Solution

From the left figure, at 100keV, the boron projected range $R_p = 0.3$ μm.

From the right figure, at 100keV, the boron straggle or standard deviation is 0.07 μm.

Chapter 9

Since $n_o = \frac{\phi}{\sqrt{2\pi}\sigma_p}$ where ϕ is the integrated dose,

$$n_o = \frac{5x10^{15}}{\sqrt{2\pi}(7x10^{-6})} = 2.85x10^{20}cm^{-3}.$$

Total number of implanted ions = area times dose = (π x 10^2) x (5 x 10^{15}) = 1.57 x 10^{18}.

Beam current = total implanted charge / implantation time = 1.6 x 10^{-19} x 1.57 x 10^{18} / 90 = 2.8 mA.

9.7 Implant Uniformity

Nonuniformity of implanted dopants across a wafer can be readily unraveled from sheet resistance measurements. ***Figure 9.12*** depicts maps obtained from three wafers implanted with 50 keV phosphorus at three different doses: 5 x 10^{13} atoms/cm^2, 2 x 10^{14} atoms/cm^2, and 2 x 10^{15} atoms/cm^2, respectively. The contour lines demarcate 1% changes in sheet resistivity. Channeling is observed in ***Figure 9.12a*** as the beam angle varies when being swept across the wafer surface. At higher dose (***b*** and ***c***), channeling is reduced because of the larger amount of incurred crystal damage.

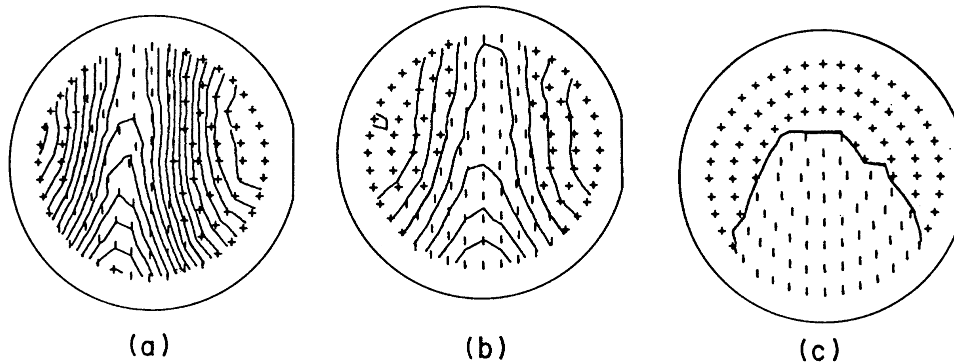


Figure 9.12: Sheet resistance maps of three samples implanted with 50 keV phosphours as a function of dose: (**a**) 5 x 10^{13} cm^{-2}, (**b**) 2 x 10^{14} cm^{-2}, (**c**) 2 x 10^{15} cm^{-2}.

9.8 Contamination

Ion implantation is basically very clean because the beam analysis separates contaminant ions from the ion beam before impacting the target. Nonetheless, there are other sources of contamination near the end of the beam line, such as adventitious metal atoms knocked from chamber walls, wafer holders, clips, masking apertures; various hydrocarbons from pump oil; cross contamination due to sputtering of old source atoms adhering on the walls; and foreign particulates dropping onto the wafer during sample loading, unloading, and transportation.

9.9 Annealing

After ion implantation, the wafer is usually so severely damaged that the electrical behavior is dominated by deep-level electron and hole traps that capture carriers and make the resistivity high. Annealing is required to repair lattice damage and put dopant atoms on substitutional sites. The success of annealing is often assessed by the fraction of dopant that is electrically active, as found experimentally using the Hall effect technique. The Hall effect measures an average effective doping level, which is an integral over local doping densities and local mobilities evaluated per unit surface area:

$$N_{Hall} = \frac{\left(\int_0^{x_j} \mu n dx\right)^2}{\int_0^{x_j} \mu^2 n dx} \qquad \textit{(Equation 9.11)}$$

where μ denotes the mobility, n is the number of carriers, and x_j is the junction depth. If the mobility is not a strong function of depth, N_{Hall} measures the total number of electrically active dopant atoms. If annealing activates all of the implanted atoms, this value should be equivalent to the dose, ϕ.

9.9.1 Furnace annealing

The annealing characteristics depend on the dopant type and the dose. There is a clear division between cases where the silicon has been amorphized and where it has been merely partially disordered. For amorphized silicon, regrowth proceeds via solid phase epitaxy (SPE). The amorphous / crystalline interface migrates toward the surface at a fixed velocity that depends on temperature, doping, and crystal orientation. Rates for undoped silicon are shown in ***Figure 9.13***. Note that the activation energy for SPE is 2.3 eV, implying that the process involves bond breaking at the interface. The presence of impurities such as O, C, N, and Ar impedes the regrowth process, as it is believed that these impurities bind to broken silicon bonds. Dopants such as B, P, and As increase the regrowth rate (by a factor of 10 for concentrations in the regime of 10^{20} atoms/cm^3), presumably because substitutional impurities weaken bonds and increase the likelihood of broken bonds.

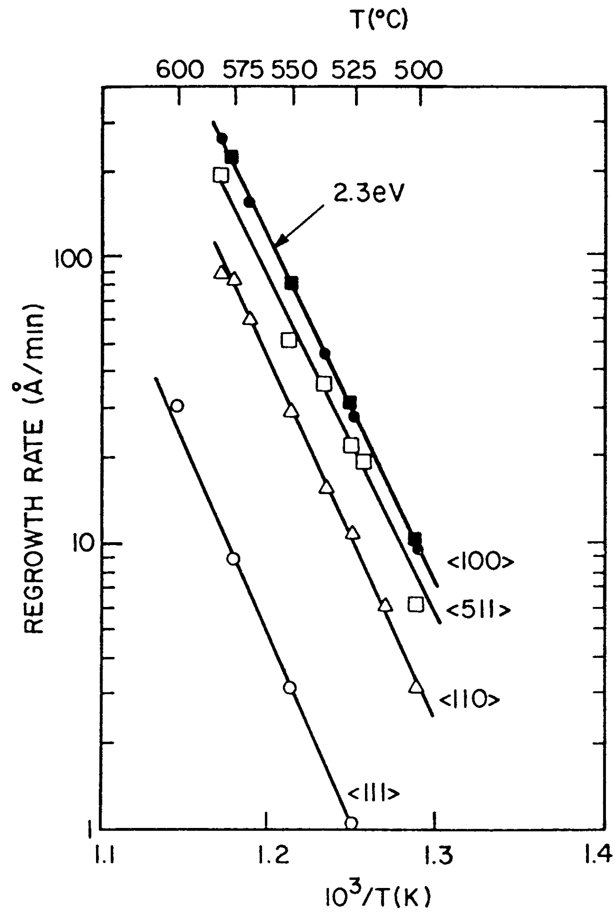


Figure 9.13: The solid-phase epitaxial growth rate of amorphous silicon as a function of temperature for various crystal orientations.

If the implantation conditions are not sufficient to create an amorphous layer, lattice repair occurs by the generation and diffusion of point defects. This process has an activation energy of about 5 eV and requires temperatures on the order of 900°C to remove all the defects. It is thus easier in many cases to repair a fully amorphized layer than a partially damaged one. The result of incomplete annealing is a reduction in the fraction of active dopant. As shown in ***Figure 9.14***, it is most severe around the peak of the dopant distribution where damage is greatest.

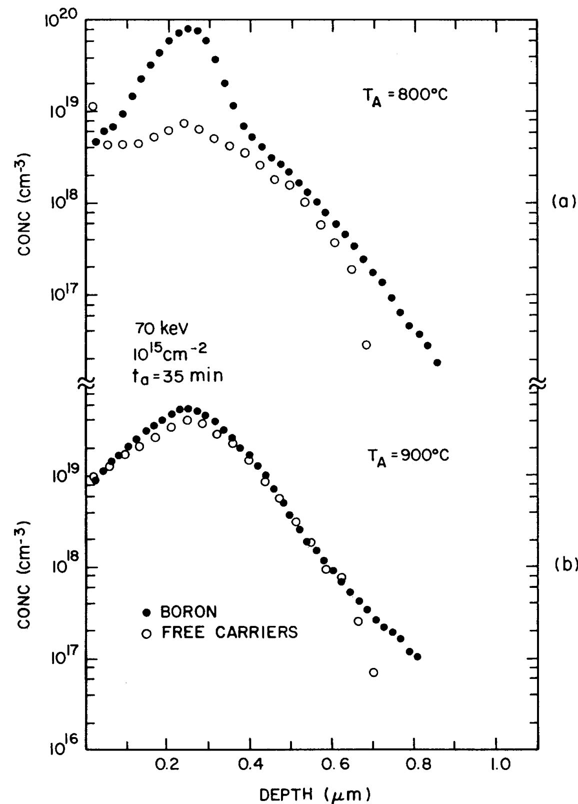


Figure 9.14: Concentration profiles of boron atoms (SIMS – solid circles) and corresponding free carrier concentrations (Hall data – open circles).

9.9.2 Rapid Thermal Annealing

The primary goal of annealing is to repair lattice damage, a process with an activation energy of 5 eV, while minimizing dopant diffusion, which typically has an activation energy of 3 to 4 eV. Because of these energy differences, at sufficiently high temperatures, repair is faster than diffusion. Furnace annealing is capable of supplying high temperature but the practical steps required to insert and remove wafers without stressing them lead to a minimum furnace annealing time of about 15 minutes. This is much longer than required to repair damages at high temperatures and so allows unnecessary diffusion. Rapid thermal annealing (RTA), which allows repair of damage with minimal diffusion, is a generic term incorporating several methods of heating wafers for periods of time from 100 seconds down to nanoseconds.

The most common type of RTA is rapid isothermal annealing using tungsten-halogen lamps to heat the wafer from one or both sides, as shown in ***Figure 9.15a***. Another method is furnace-based rapid thermal processing and the apparatus is shown schematically in ***Fig. 9.15b***. In this method, a thermal

gradient is established by adjusting the power supplied to different zones of the bell jar, with the hottest zone on top. The sample is introduced rapidly into the zone to achieve RTA. A wafer temperature up to 1200ºC and ramp rate of up to 150ºC/sec can be achieved.

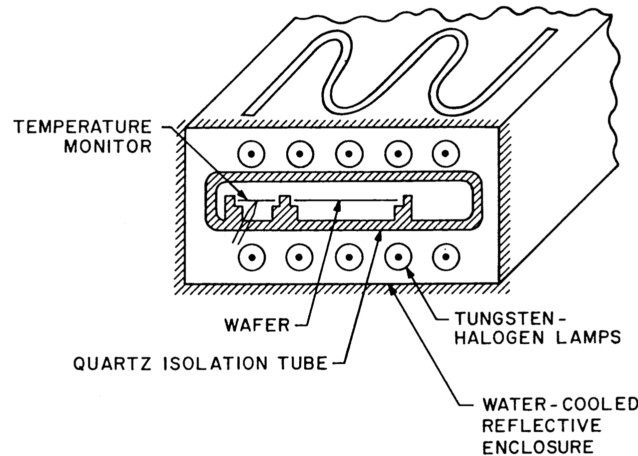


Figure 9.15a: Schematic of an isothermal rapid thermal annealing system.

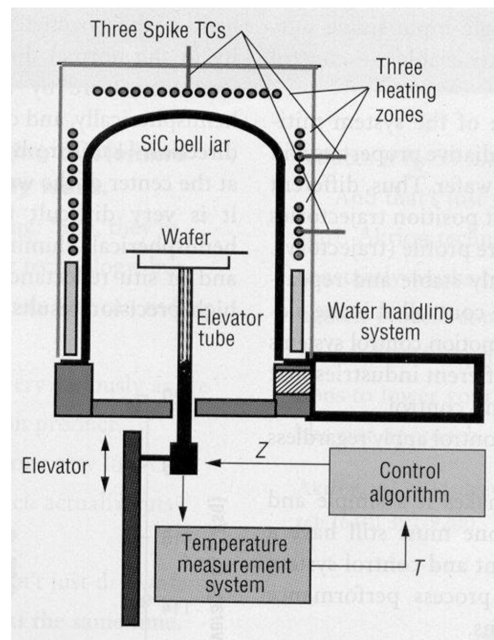


Figure 9.15b: Schematic of a furnace rapid thermal annealing system.

Both RTA schemes offer significant advantages for VLSI processing because good activation can be achieved with insignificant dopant diffusion. ***Figure***

9.16 shows an arsenic implant after RTA and almost complete electrical activation is accomplished without much diffusion. In order to fabricate ultra-shallow junctions (*discussed in more details in the next section*), spike RTA technologies employing laser heating or other means expose the wafer to a high temperature for a fraction of a second.

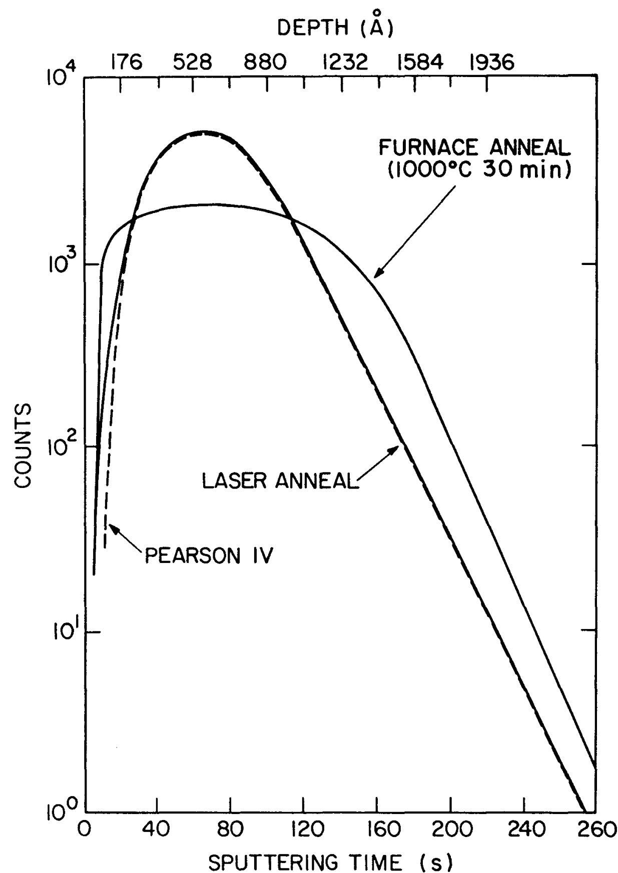


Figure 9.16: Comparison of annealed profiles using conventional furnace and laser RTA methods. Laser annealing activates the dopant without significant diffusion.

9.10 Shallow Junction Formation

In scaling horizontally to sub-micrometer and deep sub-micrometer dimensions, it is imperative also to scale dopant profiles vertically. Junction depths on the order of tens of nanometers or less are required for deep sub-micrometer devices. These can be achieved by ion implantation using various methods.

9.10.1 Low Energy Implantation

For CMOS devices, shallow N+ and P+ layers are needed for the source and drain regions. Arsenic is heavy enough to form shallow N+ layers with implantation energies permitted by commercial low energy ion implantation machines. However, in the case of boron, the implanted atoms penetrate deeper, but the effective energy can be reduced by a factor of 11/49 by implanting the molecular ion BF_2^+, because upon impact, BF_2^+ dissociates into atomic boron and fluorine. The extra fluorine atoms increase the lattice damage, thus minimizing channeling and facilitating annealing. New advances have enabled modern low energy ion implanters to deliver reasonably high ion current at energy below 1 keV. For ultra-shallow junction formation, laser doping and plasma doping are alternative techniques, but low energy beam-line ion implantation continues to be the main stream technology.

A profile can be moved closer to the surface by implanting through a surface film such as silicon dioxide. This shifts the profile by roughly the oxide thickness, but recoiled oxygen atoms can be problematic. This recoil effect can, however, be utilized if we dope by knocking dopant out of a deposited surface film using silicon or inert gas implantation, a process termed ion mixing.

9.10.2 Tilted Ion Beam

If the wafer is tilted at a large angle relative to the ion beam, the vertical projected range can be decreased, as illustrated in ***Figure 9.17***. However, for large tilt angles, a significant fraction of the implanted ions is scattered out of the surface, so the effective dose is reduced. As a pragmatic technique, this is only useful when the wafer surface is not patterned because large tilt angles cause long shadows and asymmetries at mask edges.

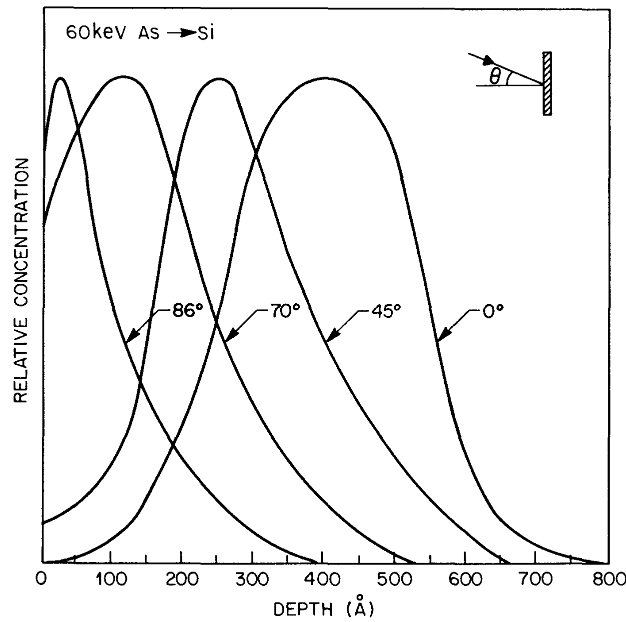


Figure 9.17: 60 keV arsenic implanted into silicon at various incident angles.

9.10.3 Implanted Silicides and Polysilicon

The problem of forming a shallow layer can be circumvented if a surface layer is deposited and the dopant is subsequently diffused into the substrate from the surface layer. This is most often done when the surface film is to be used as an ohmic contact to the substrate. Carefully controlled diffusion can result in steep dopant profiles without damaging the silicon lattice. As shown in ***Figure 9.18***, dopant diffusion in polysilicon (or silicides) is generally much faster than in single-crystal silicon, and so the implanted atoms soon become uniformly distributed in the surface thin film. Some of the dopant atoms diffuse into the substrate, consequently yielding a fairly abrupt profile like the one exhibited in ***Figure 9.18***. The small peak at the interface may be due to grain boundary segregation or impurities trapped at the interface. As shown by the dotted curve, the presence of a 25Å thick oxide layer between the polysilicon and the substrate is sufficient to block most of the diffusion. For silicides, there exists another option of implanting into the deposited metal film before the heat treatment that forms the silicide. If the implant is beneath the metal layer, the dopant atoms will be "snow ploughed" forward as the silicide forms, resulting in a steep dopant gradient near the interface. If the implant is inside the metal, it will segregate out at the moving silicide-silicon interface giving a very sharply peaked dopant distribution as well.

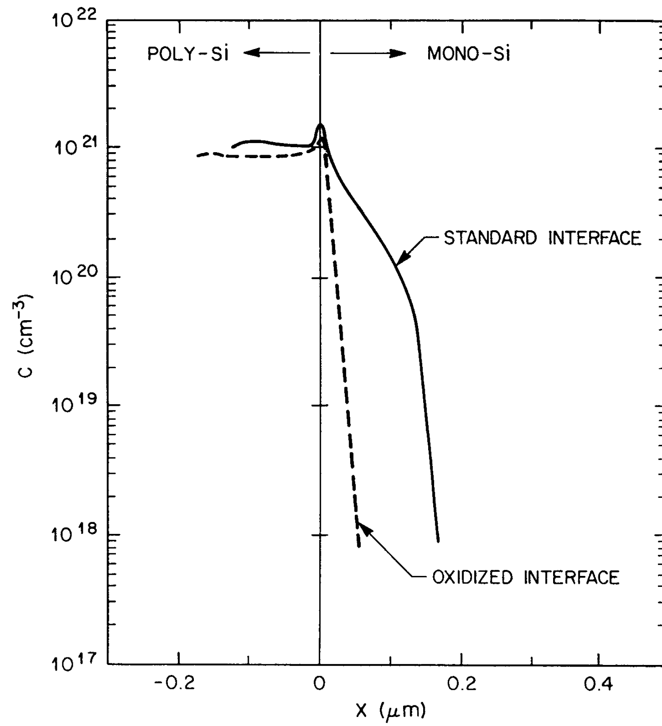


Figure 9.18: Arsenic diffusion into silicon from a polysilicon source, 950ºC for 30 minutes. The presence of a 2.5 nm oxide is sufficient to block most of the diffusion.

9.11 High Energy Implantation

Implantation at MeV energy is frequently used in VLSI fabrication. The most common application is to form deep isolation regions among individual devices. It is also used to form tubs in CMOS structures (***Figure 9.19***). High-energy ion implantation offers three ways in which the traditional epitaxial CMOS process can be improved. Firstly, the tubs can be implanted rather than diffused from the surface. In order to achieve a roughly uniform doped layer, a series of implants at different energies are required, accompanied by a short annealing step. Secondly, the structure can be improved further by retrograde doping of the tub, i.e., varying the implant doses such that the tub surface is less doped than the tub bottom. This method has the same advantages for the tub transistors as the use of epi-substrates has for the transistors outside the tub. Last but not least, the epi-substrate can be dispensed altogether by using a blanket high-energy implant in the first processing step. This will allow the formation of a buried, heavily-doped layer serving the same function as a heavily-doped substrate.

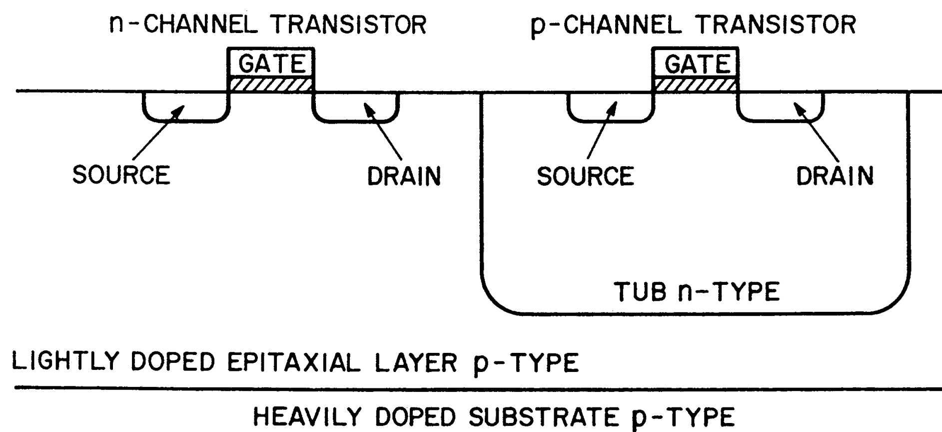


Figure 9:19: Cross section of an epi-substrate CMOS device.

9.12 Buried Insulator

Devices can be fabricated in a thin silicon layer on an insulating substrate using two types of materials: SOS (silicon-on-sapphire) and SOI (silicon-on-insulator). Both technologies have the advantage of increasing radiation hardness due to the reduced collection volume for charges generated by ionizing radiation. They also offer a compact way to isolate devices from each other to reduce parasitic capacitance and to eliminate latchup for CMOS circuits.

The SOS technology is more mature. High quality silicon epitaxial films have been successfully grown on sapphire wafers. However, sapphire substrates are very expensive, thereby limiting the use of SOS to demanding applications, such as military devices.

One of the two common SOI techniques is to introduce a blanket buried oxide to isolate the device active region from the bulk wafer, a process called SIMOX (separation by implantation of oxygen). The principle behind the formation of a buried oxide layer is quite simple. If oxygen is implanted at a dose on the order of 10^{18} atoms/cm^2, there will be twice as many oxygen atoms as silicon atoms at the vicinity of the peak. Upon annealing, silicon dioxide will form.

The details of high dose oxygen implantation differ from traditional implantation steps in several ways. The goal is to maintain a surface layer of high-quality single-crystal silicon for device fabrication. Hence, the substrate is kept near 600°C during implantation so that self-annealing maintains the crystal integrity. Each incident oxygen ion sputters on the order of 0.1 silicon atom, and so the large number of implanted oxygen ions erode many layers of silicon atoms. Nonetheless, this is more than offset by the expansion in volume during the formation of oxide (44%). The net result is a slight swelling of the silicon surface. The implant profile also changes from a Gaussian-like distribution to a flat-topped distribution, as depicted in ***Figure 9.20***. This is due to diffusion of oxygen to the silicon-oxide boundaries after oxygen saturates the substrate around the peak region. As implanted, the surface layer still contains a substantial amount of oxygen and much damage, albeit still single-crystal. Annealing is performed at a high temperature above 1300°C to cause a strong segregation of oxygen into the buried layer from both sides, consequently depleting the surface of almost all the implanted atoms, including impurities. This leaves a high-quality surface film containing very little oxygen and less than 10^9 dislocations per cm^2.

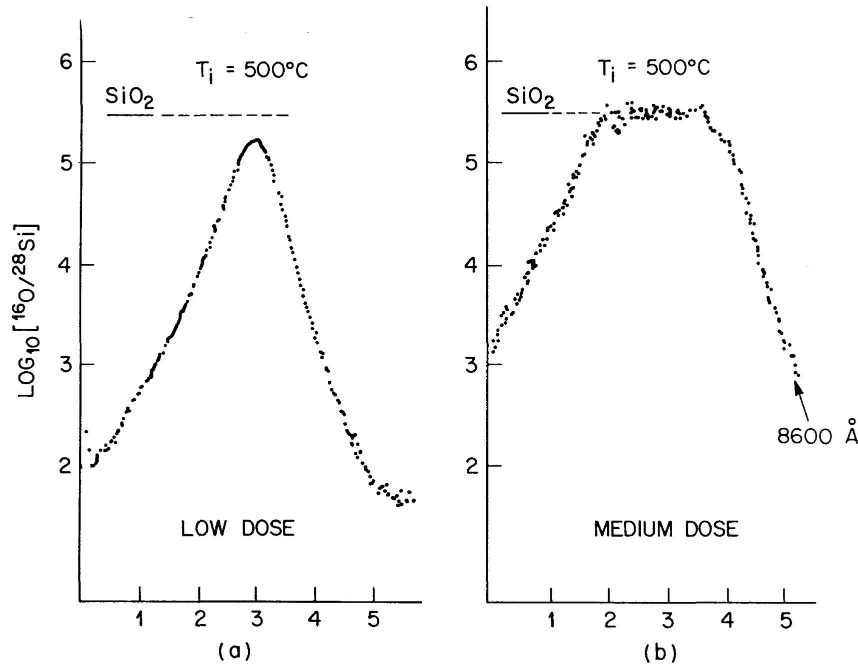


Figure 9.20: Unannealed profiles for 200 keV oxygen implanted into silicon. (**a**) For a low oxygen dose, the profile has the usual Gaussian shape. (**b**) For a high enough oxygen dose and after high temperature annealing, oxygen diffuses toward the implant peak to form a stoichiometric buried silicon dioxide layer.

SOI substrates can be synthesized using wafer bonding and back etching. Using wafer bonding technology, two silicon wafers (one or both having a surface oxide layer) can be pasted together resulting in a Si / SiO_2 / Si structure. One side of the structure can be thinned by polishing to yield the required silicon layer thickness. The advantage of this technique is that the thickness of the oxide and silicon layers can be adjusted independently, but two wafers are required to make one SOI wafer thereby raising the cost.