Defects in Semiconductors: Some Fatal, Some Vital

Hans J. Queisser* and Eugene E. Haller

REVIEW

The role of defects as essential entities in semiconductor materials is reviewed. Early experiments with semiconductors were hampered by the extreme sensitivity of the electronic properties to minute concentrations of impurities. Semiconductors were viewed as a family of solids with irreproducible properties. Scientific efforts overcame this idiosyncrasy and turned the art of impurity doping into today's exceedingly useful and reproducible technology that is used to control precisely electrical conductivity, composition, and minority-carrier lifetimes over wide ranges. Native defects such as vacancies and self-interstitials control basic processes, foremost self- and dopant diffusion. The structural properties of dislocations and higher dimensional defects have been studied with atomic resolution, but a thorough theoretical understanding of their electronic properties is incomplete. Reactions between defects within the host lattices are increasingly better understood and are used for gettering and electrical passivation of unwanted impurities. Metastable defects such as DX centers and the EL2-related arsenic antisite are briefly discussed. The recent development of isotopically controlled semiconductors has created new research opportunities in this field.

The perfect semiconductor single crystal is dull compared with the vast variability of today's deliberately engineered materials. The ideal model for a solid demands uninterrupted translational symmetry. Every atom resides on its prescribed site; no impurity or defect is permitted. Abrupt symmetry terminations at surfaces are ignored in this idealization, which then enables theorists to derive the essential intrinsic properties, such as the energies of the valence and conduction bands of a material. "Imperfections in a nearly perfect crystal" is the guiding principle for semiconductor science and technology. The electronic band structure, which relates electron energy E to its wave vector $\mathbf{k}$, or the spectrum of lattice vibrations can be calculated. Group theory explains some basic features such as piezoelectricity or the selection rules for optical transitions between the energy bands of the solid. Real semiconductor materials belie this ideal (1).

Before the 1950s, experiments with semiconducting natural minerals (like galena, PbS) gave puzzling results. Parameters such as electrical conductivity or optical absorption were scattered wildly. Physicists hesitated to admit such materials with irreproducible properties into their realm of quantitative investigations. The explanation for these results was the strong fluctuation of composition and purity in these minerals. Although most solids maintain their essential properties upon introduction of small amounts of foreign atoms or when their lattices do not perfectly extend over large distances, certain semiconductor properties—especially conductivity, free carrier mobility, and carrier lifetime—change dramatically with the introduction of defects and impurities at the parts per million or parts per billion

Department of Materials Science and Mineral Engineering, University of California and Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

*Miller Institute Visiting Professor, on leave from Max-Planck-Institut für Festkörperforschung, D-70506 Stuttgart, Germany.

level. A new technical strategy for the preparation of semiconductors arose with the discovery of the transistor in 1947. It was recognized that pure and structurally perfect crystals had to be prepared with effective purification and crystallization methods. These highly perfect crystals were then doped to obtain the desired electronic functions inside the crystal. Today, typical doping levels for silicon devices range from 10^{13} to about 10^{20} dopant atoms per cubic centimeter, or approximately one impurity atom per 10^{10} to about 10^3 host atoms. The word "defect" may be scientifically correct, yet it does not convey the tremendous utility of these entities. This extrinsic behavior has been thoroughly investigated during the past 50 years and has led to the creation of the multitude of applications of today's semiconductor materials (2).

It is not possible to cover in this brief article the immense variety of defects in semiconductors and their experimental and theoretical studies. Our particular choice of topics does not reflect on the importance of those issues that are not discussed.

Shallow Versus Deep Defects

The band gap, the energy region that is free of electronic states and is bounded by the valence and conduction bands, is the hallmark that differentiates metals, semiconductors, and insulators. Dopants and defects can dramatically modify the electrical conductivity of semiconductors as they add states in this band gap, which can provide electrical charge carriers (electrons or holes) to the respective bands at energies very much lower than the gap energy. Electrically active defects in semiconductor crystals have different characteristics, depending on the location of their energy states relative to the conduction- or valence-band edge. Shallow defects have energy levels within a few tens of millielectron volts from the respective band edges, whereas deep defects typically reside within the middle third of the semiconductor energy band gap. Today this simple definition is no longer generally applicable. Deep levels have highly localized wave functions whereas shallow-level wave functions are as extended as the far reaching Coulomb potential. According to this definition, oxygen in GaN is a deep defect even though it is located inside the conduction band (3).

A shallow defect is easily explained. For example, a shallow donor resembles a hydrogen atom with a positive nucleus binding an electron. Two modifications to the hydrogenic model arise inside a solid. First, the mass of the electron is renormalized because of its interactions with the lattice. This "effective mass" is usually smaller than that of the free electron. Second, the crystal reduces the binding potential, as described with a dielectric constant. Combined, these effects reduce the ground-state binding energy (ionization energy) of a donor, from the well-known 13.6 eV of the isolated hydrogen atom, down to merely 45 meV in Si. Other hydrogenic features remain: a series of bound excited states, which are observable with far-infrared spectroscopy, and the ionization into the energy continuum, that is, the conduction band in the case of a donor. This strikingly simple and effective hydrogenic model of shallow-level impurities or dopants has been amply confirmed with a broad range of electrical and optical studies of a vast number of dopants in numerous semiconductor materials (4).

The deep defects are much more difficult to understand (5). Badly fitting impurity atoms distort the host lattice so extensively that the

binding energy greatly exceeds that of shallow, hydrogenic defects. The energy level lies "deep" within the forbidden energy band gap. Despite the availability of highly sophisticated computational tools, deep defects still represent one of the most difficult family of problems of current semiconductor theory. Many atoms around the distorting defect must be considered when a first-principles calculation is attempted; the defect extends over many lattice constants. The carriers interact strongly with the lattice. The effects of deep defects are to drastically reduce minority carrier lifetimes and to act as traps for charge carriers. Some deep defects can be used to pin the Fermi level near the middle of the energy gap and thus lead to highly resistive materials. The EL2 defect is widely used for growing semi-insulating GaAs crystals.
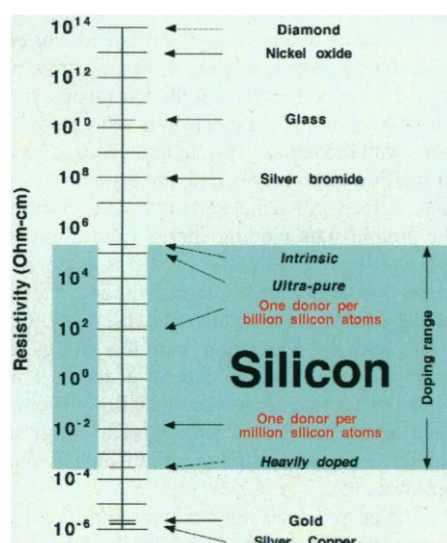
Doping to Control Electrical Conductivity

Doping a semiconductor with foreign atoms (impurities or dopants) can alter the electrical conductivity by many orders of magnitude. A phosphorus atom with five valence electrons is a donor of an electron to the conduction band when it is used to replace a silicon host atom that has four valence electrons; thus, the conduction by electrons is enhanced. An atom from group III of the periodic table, on the other hand, accepts an electron to complete its Lewis octet in the silicon lattice and thus introduces conduction by holes, each displaying a positive charge. Many, but not all, semiconductors can therefore be doped for n-type or electron conduction as well as for p-type or hole conduction. The range of doping for Si, our technologically most important semiconductor, is shown in Fig. 1. The name semiconductor is justified by this vast variation, ranging from metallic conduction to essentially insulation behavior.

The elemental semiconductors Si and Ge are easily doped for both conductivity types. These elements are thus beautifully suitable for the fabrication of bipolar devices with both n- and p-type regions, such as transistors, solar cells, or junction diodes. The essential physical phenomena occur at the interfaces between the n-type and p-type portions of the material. These p-n junctions are strongly nonlinear in current transport. The result is a rectifier with a large current flowing when a positive voltage is applied to the p-type portion and very little current crossing the junction for the opposite polarity. The first junction transistors made 50 years ago demonstrated superiority over initial devices with metallic contact needles (6).

The ability to precisely control the concentration of donors and acceptors inside a single crystal is essential for the integration of electronic functions on a semiconductor wafer. The fabrication of a modern integrated circuit (for example, a processor chip) assembles more than 100 million device functions on one chip, an approach completely different from the preceding vacuum tube technology where devices were built out of individual components. The technical and commercial success to continually miniaturize microelectronic circuits depends on this principle of integration and defined doping. Many doping methods are used: for example, donors and acceptors can be diffused from the surface inward at elevated temperatures. In another method bombardment of the lattice with specific energetic ions from an accelerator (ion implantation) circumvents the high diffusion temperatures and long times and can achieve intricate doping profiles. Thermal annealing of the radiation damage is a much more benign process than diffusion at high temperatures.

Optoelectronic devices, such as light-emitting diodes or p-n junction lasers, consist of compound semiconductors, such as gallium arsenide, gallium phosphide, indium phosphide, and numerous ternary (such as $Al_xGa_{1-x}As$ and $In_xGa_{1-x}As$) and quaternary (such as $In_xGa_{1-x}As_yP_{1-y}$) alloys. Their electronic band structure favors direct and efficient coupling of the crystal's electrons and holes to electromagnetic radiation. These compounds often resist bipolar doping. For example, gallium nitride (GaN) is well suited with its large forbidden energy gap for the emission of blue light when a bias is applied in the conduction direction to a p-n GaN junction diode. Donor or n-type doping is easily achieved, but the opposite—efficient p-type doping—cannot be easily attained. There are two reasons for this: first, the shallowest acceptors are Mg impurities on Ga sites with energy levels higher than 160 meV above the valence band top. Only about 1% of these strongly bound holes reside in the valence band at room temperature. Second, the maximum Mg concentration that can be introduced in GaN is between 10^{18} to 10^{19} cm^{-3}. Together, these two properties result in relatively small free-hole concentrations. Much current research effort is therefore directed at an understanding of these asymmetries of doping and at searching for useful impurities and intelligent ways of introducing these impurities without damaging the lattice. A review in this issue by S. Nakamura (7), one of the pioneers of the group III nitrides, is devoted to the exciting developments in the nitride semiconductors. One result of recent theoretical calculations (8) for the nitrogen vacancy, a native defect for a while thought to be the source of unintentional n-type doping, is shown in Fig. 2. These theoretical results, however, conclusively showed that nitrogen vacancies require energies that are too high for normal growth and processing conditions to form, making them unlikely candidates for the donors found in As-grown nitride semiconductors.

Doping to Control Carrier Lifetimes

When electrons cross the boundary from n-type to p-type or holes from p-type to n-type material, they change from being majority to minority carriers. The transistor inventors made the important, seren-

Fig. 1. Range of electrical resistivities of pure and donor-doped silicon single crystals shown in comparison with metals and insulators.

Fig. 2. Atomic geometry and charge density contour surface of a nitrogen vacancy in GaN. Gallium atoms are green, nitrogen is blue, and the defect wave function is shown in pink. The strong interaction of the Ga dangling bonds pushes the donor level up into the conduction band; reprinted with permission from (8).
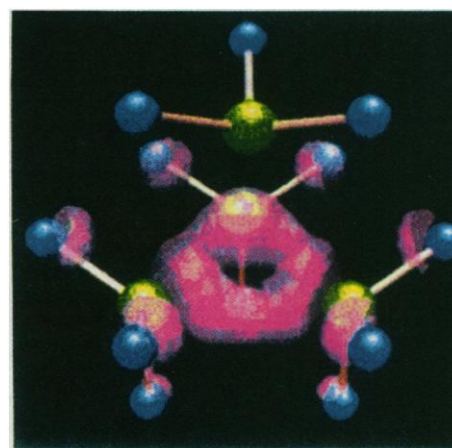
dipitous discovery that minority carriers can exist for surprisingly long times before they recombine with an oppositely charged carrier or are trapped at a defect site. A minority electron lives long enough to traverse the p-type transistor base layer on its way from the n-type emitter to the n-type collector. Such long lifetimes, however, demand that materials be perfect. Deep defects are the most effective lifetime killers. Minority carrier lifetimes are easily measured, for example, by monitoring the excess conductivity caused by a short flash of light onto the semiconductor. Lifetime became the incorruptible figure of merit for defect reduction in crystals.

Just as for reliable control of doping, reliable control of carrier lifetime is achieved by first growing a pure and structurally perfect crystal, and then introducing lifetime-determining defects. In the 1940s and 1950s, many researchers were concerned that the preparation of single crystals would be too complicated and energy consuming to be acceptable for electronics manufacturing. This apprehension was wrong. The modern semiconductor chip industry will soon use highly perfected and homogenized silicon wafers cut from crystals with diameters up to 300 mm (9). Amorphous or polycrystalline semiconductors would have severely hampered transistor development. Such imperfect materials have become, however, quite useful for inexpensive, large-area devices (solar cells and displays) where electronic performance standards can be lowered.

Very short carrier lifetimes are not always detrimental. They are needed for fast switching diodes and other devices for high-frequency operation. Well-defined techniques exist for doping with foreign atoms that practically eliminate carrier lifetime (such as gold in Si), or with point defects, judiciously introduced by irradiation with high-energy particles. The first gold-doped silicon diodes switched faster than could be displayed by the oscilloscopes of that time.

Dislocations: Line Defects in Crystals

Imagine cutting a crystal halfway and inserting an extra plane of atoms. This half-plane terminates in a line defect, the "dislocation." Early efforts at growing semiconductor crystals led to large concentrations of dislocations that arose from the intense thermal gradients generated when the crystal was pulled from the melt. There are various types of dislocations; dislocation dissociation generates two-dimensional stacking faults (10). The electronics community feared that dislocations might short-circuit narrow base layers of bipolar

transistors. Great efforts succeeded to produce dislocation-free Si (11). Thermal processing, however, reintroduces new dislocations. Fortunately, dislocations are not as harmful as expected, but their interplay with metal impurities is deleterious in electronic devices (11).

Most optoelectronic devices malfunction with dislocations present because these defects cause rapid recombination of holes with electrons without conversion of their available energy into photons. Nonradiative recombination arises, uselessly heating the crystal. Identification and eradication of dislocations was the prime task before the now omnipresent GaAs-based light-emitting and laser diodes achieved their long operating lifetimes. Currently this history repeats itself in the blue spectral regime, but in an unexpected variant. Gallium nitride, with properties closer to a ceramic material than a proper semiconductor, has become the optimal choice for blue laser-light (12). The dislocation density in this material has been huge— more than 10^9 to 10^{11} cm^{-2}, which is unprecedented—yet, efficient light-emitting diodes and lasers are feasible in these heavily dislocated crystals. This lucky surprise still awaits a scientific explanation. One spectacular example of crystal growth at a single screw dislocation on the nitrogen face of a GaN surface is shown in Fig. 3 (13).

Defect Elimination: Gettering and Passivation

Semiconductors undergo rough treatments in chip manufacturing. Ions are shot into the lattice at high speeds, oxidation and diffusion processes are performed at temperatures often beyond 1000°C in very clean but not absolutely sterile environments, plasma etching induces damage, and every treatment with a wet chemical emits massive amounts of fast-diffusing hydrogen into the lattice. These defects, above all the metallic impurities (such as iron from the furnace wiring), must be rendered harmless (14). There are two common remedies: gettering the dirt into an unused portion of the material, and passivating the electrically active but undesirable impurities.

Gettering is an old, highly effective remedy of the early vacuum-tube age (and is still used in cathode ray tube displays). Evaporation of a pill of magnesium or barium onto the inside of the vacuum tube glass enclosure after thorough evacuation greatly improved the vacuum inside the bulb by binding residual nitrogen and oxygen that outgassed during tube operation. The same phenomenon is now applied throughout the chip industry. Metallic impurities are gettered by silicon dioxide (SiO_2) precipitates. Carefully devised thermal treatments generate the SiO_2 inclusions, which in turn efficiently bind most metallic impurities. Intricate thermal processes generate such precipitation at a safe distance from those near-surface regions where the active parts of a modern device are located (Fig. 4) (15, 16). Chip makers, therefore, use silicon crystals with a tightly specified oxygen content. Such crystals are grown out of a silica crucible, which serves as the oxygen source. Oxygen is usually not electrically active in Si, although under special conditions oxygen may form complicated multi-atom centers, called "thermal donors," which are avoided in device processing by annealing procedures. This complex set of

Fig. 3. Scanning tunneling micrograph of a single screw dislocation on a GaN crystal surface of wurtzite structure. Two Ga-N bilayer growth fronts surround the dislocation, and several reconstructed surface domains are clearly visible; reprinted with permission from (13).

Fig. 4. Cross-sectional view of the defect-free, near-surface region of a silicon wafer. The lower portion of the figure shows silicon dioxide precipitates used for impurity gettering; reprinted with permission from (15).

20μm
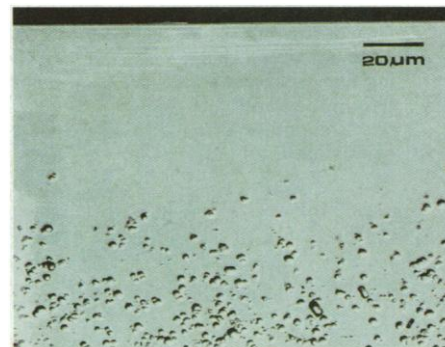
process steps is a marvelous example of the meticulous control of defects within Si.

Some unwanted electrically active centers in a semiconductor can be passivated (or neutralized) by the formation of neutral pairs of the impurities with interstitial lithium or hydrogen. The technique of drifting Li^+ ions in the electric field of a reverse-biased diode had been used for many years to passivate acceptors in p-type Si and Ge. The process leads to millimeter- or centimeter-wide depletion layers that are widely used in solid-state ionization detectors for x-rays, gamma rays, and energetic particles (17). The development of ultra-pure Ge with one dopant impurity per 10^{12} Ge atoms (a tour de force resulting in an electrically "twelve-nines" pure substance) has made Li^+ drifting in this semiconductor obsolete and has made possible the fabrication of Ge detector diodes with depletion zone volumes of several hundred cubic centimeters (18). All modern high-energy resolution x-ray detectors used, for example, on transmission and scanning electron microscopes are fabricated from Li^+ ion–drifted p-type Si. The resolution of these devices is superior to those made from high-purity Si because Li^+ ions passivate shallow as well as deep-level dopants and defects (19).

During the development of ultrapure Ge, hydrogen was discovered to be an impurity that could either activate or passivate impurities and defects. For example, oxygen, silicon, and carbon were found to form shallow-dopant complexes with hydrogen (20). Multivalent acceptors and deep-level impurities became partially or fully passivated by combining with hydrogen (21). In the early 1980s, hydrogen was found to enter Si and to passivate acceptors and donors (22). Almost every device processing step that involves wet chemicals (even boiling in water) will introduce hydrogen into Si. Deliberate introduction of hydrogen with proton implantation is used to form highly resistive layers to isolate neighboring devices or to confine electrical currents to specific areas. Separating a thin layer from an underlying substrate can be achieved with implantation of large doses of hydrogen. An entire new body of solid-state chemistry describing hydrogen-impurity and hydrogen-defect complexes in Ge, Si, and many compound semiconductors has been created in recent years. One example of the multiatom center involving platinum binding two hydrogen atoms is shown in Fig. 5 (23). Molecule formation and dissociation, reaction kinetics, and thermodynamic properties, such as the entropy contribution of a complex inside the semiconductor, are aspects of this rapidly expanding subfield of defect science.

Semiconductors with Isotopically Controlled Composition

Pomeranchuk suggested in 1942 that perfect crystal symmetry is already perturbed when the constituent atoms vary in isotopic atomic mass and that this isotope disorder should reduce the thermal conductivity (24). The isotope separation efforts during and after World War II led to sizable quantities of enriched stable isotopes of numerous elements, among them isotopically pure ^{74}Ge. Geballe and Hull finally proved the conjecture of Pomeranchuk in 1958 (25). The enormous expense associated with isotopically enriched materials (tens of U.S. dollars per milligram) stifled further studies until the end
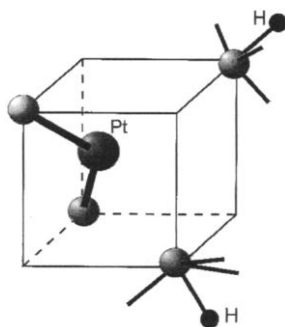
of the cold war when collaborations between former Soviet and Western scientists made research with isotopically controlled semiconductors feasible. During the past few years, defect studies with these special materials have generated a host of results (26), two of which will have to suffice for illustrating the possibilities.

Oxygen in Si or Ge is not an electrically active dopant impurity. It plays, however, a crucial role in gettering transition metal ions in Si, in forming shallow donors with H in Ge, and in other processes. Oxygen assumes a bond-centered interstitial position in these materials. In that position it is infrared active, forming sharp absorption lines near 862 cm^{-1} in Ge (8.06 cm^{-1} is equivalent to 1 meV). Because natural Ge consists of five stable isotopes with sizable concentrations, the oxygen atoms have a variety of nearest neighbor atom combinations, each one leading to slightly different vibrational frequencies. The resulting absorption spectrum shows 11 distinct lines (Fig. 6A). The close proximity of these lines forbids any detailed studies of line splittings caused by the interaction of the various possible vibrational modes (v_1, v_2, and v_3). Growth of isotopically pure, oxygen-doped Ge crystals drastically simplified the spectrum to just one line (Fig. 6B) and allowed the determination of the interactions between the v_1, v_2, and v_3 modes (27).

Self-diffusion is the most fundamental, atomic, random walk process in a solid. In semiconductors it is the starting point for the understanding of all dopant diffusion processes. Radioactive isotopes have typically been used to follow the random walk of host atoms. Both the short half-life of the most suitable radio isotope (^{31}Si with a half-life of 2.6 hours) and the near-surface effects have made an accurate determination of the Si self-diffusion coefficient rather tentative. Epitaxial growth of a multilayer structure consisting of natural and ^{28}Si together with secondary ion mass spectrometry (SIMS) enabled the first direct determination of Si self-diffusion profiles (28). A single diffusion enthalpy of 4.75 ± 0.04 eV characterizes the self-diffusion process within these limits. This result is attractively simple and quite unexpected. It will force revisions of dopant diffusion simulation and modeling of advanced microchips. Diffusion of dopant impurities is controlled by native defects. Most donor and acceptor dopants diffuse in Si assisted by Si interstitials, with Sb being the notable exception. Any process that significantly changes the native defect concentration will affect dopant diffusion. Oxidation and
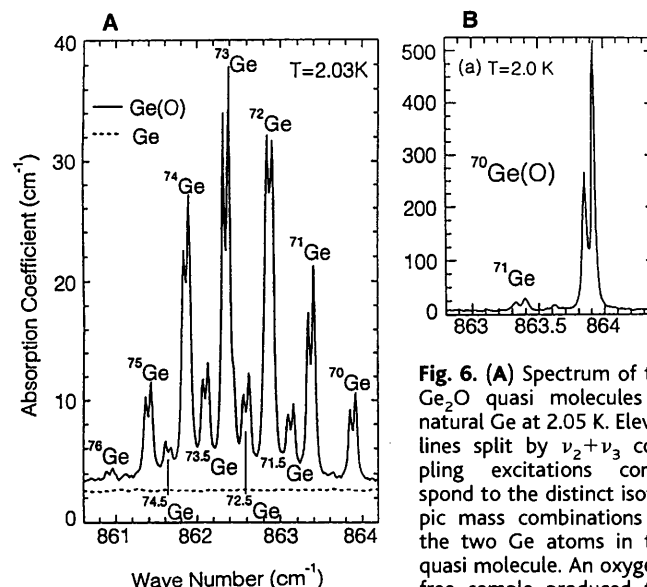
Fig. 5. Model of a platinum atom binding two hydrogen atoms in the silicon lattice; reprinted with permission from (23).

Fig. 6. (A) Spectrum of the Ge_2O quasi molecules in natural Ge at 2.05 K. Eleven lines split by v_2+v_3 coupling excitations correspond to the distinct isotopic mass combinations of the two Ge atoms in the quasi molecule. An oxygen-free sample produced the dashed line. **(B)** The spectrum of an oxygen-doped, highly enriched sample of ^{70}Ge. The lines labeled ^{71}Ge are caused by traces of ^{72}Ge in ^{70}Ge; reprinted with permission from (27).

very large donor concentrations increase the Si interstitial concentration, leading to an interstitial "wind" moving from the surface toward the interior. This wind causes transient enhanced dopant diffusion, one of the most serious problems for modern devices (29).

Surfaces and Interfaces

Every semiconductor is bounded by surfaces, and most modern devices also have a number of internal interfaces: homointerfaces between equal materials and heterointerfaces between different kinds of materials. Intuition suggests that these locations should be the sources of very large concentrations of a variety of structural defects that would have large effects on electrical, optical, and other properties. Although this expectation is fulfilled for a majority of cases, the most important interface formed between crystalline Si and vitreous SiO_2 has an extremely low density of electrically active states. Only one in 10^3 to 10^4 interface bonds of a properly prepared oxide leads to an energy level in the band gap of the Si. This hallmark property of Si has made possible the metal-oxide-semiconductor (MOS) technology used in the fabrication of the vast majority of modern Si integrated circuits.

In contrast, metal-silicon interfaces are most often dominated by interfacial defects, with a distribution of energy states peaking near one-third of the band gap energy above the top of the valence band. Under these conditions, the property of a metal-silicon junction, also called the Schottky barrier, is only weakly dependent on the type of metal used. Reaction between certain metals and silicon leads to the formation of silicides. A large number of studies of such silicide-silicon interfaces show that the spectrum of interfacial defects changes drastically, leading to Schottky barriers that are strongly dependent on the particular silicide (30). Silicides are the preferred Ohmic contact materials for many device applications. In these cases, the Schottky barrier is short-circuited by tunneling of charge carriers between the metallic silicide and a heavily doped, near-surface layer in the semiconductor.

Metal-to-GaAs interfaces present an especially important case. Again, such barriers are independent of the particular metal, and it has been proposed that native defects, in this case arsenic atoms on gallium sites (so-called As-antisites) dominate the interface properties. A different theoretical model is based on metal-induced gap states. Neither model can explain the large number of experimental findings. The current thinking favors a model that uses both donor- and acceptor-type native defects, which partially compensate one another and stabilize the Fermi level at a specific energy (31).

Surfaces, which are most often covered with oxides, and the metal-semiconductor interfaces differ from today's variety of semiconductor-semiconductor interfaces. The defect density at such interfaces is most carefully controlled through the choice of compound and alloy semiconductors with closely matched crystal lattice constants, elaborate surface cleaning procedures, and sophisticated thin-film epitaxial growth techniques. With all of these measures, both homo- and heterointerfaces can be grown practically defect-free. This mature technology is widely used in the growth of multilayer structures of $Al_xGa_{1-x}As$ and GaAs, doped n- and p-type, for the fabrication of double heterojunction lasers. Today, every compact disc (CD) player uses such a diode laser for reading the information imprinted on the CD.

The recent progress with the group III–nitride semiconductors, which has led to bright blue light–emitting diodes and lasers, has refocused attention on interfaces. At this time, GaN single crystals cannot be grown in bulk form of sufficient size for device fabrication. Therefore, all GaN-based devices consist of thin films grown either on sapphire (Al_2O_3) or on silicon carbide (SiC). Large lattice mismatches exist in both cases, and until very recently, epitaxial growth efforts have led to extremely large dislocation densities. Despite these large defect densities, efficient light-emitting devices can be fabricated. Growth through narrow openings of SiO_2 masks deposited on highly defective GaN layers followed by lateral overgrowth on the mask material has led to vastly improved epitaxial layer quality. This rapid progress is another example of the ingenuity of the scientists and engineers who always find ways to control defects when required.

Metastability: DX Centers and the EL2 Defect

The optoelectronic properties of group III–V compound semiconductors, especially GaAs and its alloys with AlAs and GaP, began to attract attention in the 1960s and 1970s. With alloying it became possible to change the band gap and the color of light-emitting diodes and lasers continuously over a wide range of the visible and infrared spectrum.

The lattice-matched AlAs-GaAs system was especially attractive because a variety of superlattice and quantum well structures could be grown epitaxially and defect-free on high-quality GaAs substrates. However, one serious drawback appeared when n-type doping was attempted with $Al_xGa_{1-x}As$ with an AlAs mole fraction $x \geq 0.22$ (32). The standard shallow donors (for example, Si on Ga sites or S on As sites) that provided free electrons for various device functions were found to turn into deep-level defects. This conversion to deep centers put devices that required layers with large AlAs mole fractions out of reach. Lang and Logan (33) coined the expression DX centers, assuming that the donors combined with an unknown defect X, forming the deep-level centers. At that time the possibility was not considered that a defect could change its geometric configuration, that is, that it could have metastable configurations or could even be bistable. An additional puzzling observation was the persistent photoconductivity (PPC), which is another misnomer. Alloy crystals that contained DX centers and were cooled to very low temperatures would conduct during illumination with photons of energy larger than the band gap and surprisingly would continue conducting electricity after the light was turned off.

An interesting discovery was made by Mizuta et al. (34) when they exposed small n-type GaAs samples to large hydrostatic pressures and found them to lose their conductivity at pressures around 20 kbar. This result was the first strong indication that it was not the AlAs per se that leads to DX center formation but rather the change in band structure. The changes caused by hydrostatic pressure or by alloying with AlAs are similar in GaAs. (A 1% AlAs mole fraction changes the band structure in a similar way to the application of a 1-kbar hydrostatic pressure.) The pressure experiments established that single isolated shallow donors convert to DX centers. From this point on, X only signified a different configuration and no longer a specific defect needed to combine with a shallow donor to form DX centers.

A further experimental observation was the large difference in the optical and the thermal ionization energy of DX centers. Commonly known as large Stokes shift, this observation supported a large lattice relaxation model for the shallow-donor to DX conversion. All of these unusual experimental findings found their explanations in the model proposed by Chadi and Chang (35). Their calculations showed that there exist two geometric configurations for Si (or other group IV elements) on or near a Ga lattice site and S (or other group VI elements) on or near an As lattice site. The shallow-donor configuration is the tetrahedrally symmetric configuration with four identical chemical bonds, one with each host lattice atom neighbor. The DX configuration for group IV donors reduces the number of bonds to three and moves the donor atom away from one As neighbor by more than 1 Å. An additional twist is the necessity of DX to bind an extra electron, which stabilizes the DX configuration. The bonding configuration of the As atom that lost its bond to the Si donor is quite common. For example, AsH_3 has three bonds and one lone electron pair.

The configuration coordinate (CC) diagram provides a powerful framework (36) for interpreting Chadi and Chang's model. The total defect energy is plotted as a function of a generalized coordinate, which is an overall measure of all the changes in the defect atom and

its neighbor's coordinates. The CC diagram for the DX center described by the model of Chadi and Chang is shown in Fig. 7. At low pressures or AlAs mole fractions, the CC diagram to the left applies. The shallow-donor configuration is stable, but DX coexists in metastable form. Optical excitation can lead to DX formation. At modest temperatures, the small energy barrier separating the two defect structures is easily overcome, returning the center into its stable form. At pressures $\geq$20 kbar or at AlAs mole fractions $\geq$22%, the DX configuration is stable. The material is insulating despite the presence of dopants. During band edge light illumination free electrons and holes are generated, leading to the standard photoconductivity. However, DX centers are also optically pumped into the shallow-donor configuration. If the crystal is kept cold, the shallow-donor configuration is trapped by the energy barrier between the two configurations. The donors are very shallow and remain ionized, that is, the conductivity persists, no longer because of photon illumination but because of the photon-induced defect configuration change.

Although the Chadi and Chang model was broadly accepted, the requirement for DX to be negatively charged was questioned. If true, at most half of all the donors (d°, d^+) could be converted to DX centers: $2d^\circ \leftrightarrow d^+ + DX^-$. Wolk et al. (37) found an elegant way to prove experimentally that the DX centers are indeed negatively charged. Combining large hydrostatic pressure with far-infrared absorption spectroscopy, they showed that the vibrational mode frequencies of the tetrahedral shallow Si donor and the trigonal Si DX center were different and that the absorption line strength of the shallow donor was at least as large or larger than the DX line strength. Using n-type GaAs samples with different acceptor concentrations, they showed unambiguously that DX centers are stabilized by an extra electron.

A second metastable defect is associated with a deep donor (EL2) found in GaAs crystals with excess As concentrations. The surplus As atoms were found to assume Ga host lattice sites, forming antisites, As_{Ga}. For many years it was unclear if the isolated As_{Ga} antisite or a more complicated defect is responsible for forming the EL2 donor level. This debate was settled by the theoretical work of Dabrowski and Scheffler (38) and Chadi and Chang (39) in favor of the isolated antisite.

The EL2 donor exhibits fascinating properties that show similarities to those of the DX center. For example, EL2 can be optically pumped with photons between 1 and 1.3 eV into a metastable state with an identical charge state but no energy level in the band gap. From this electrically and optically inaccessible metastable state, only a return with thermal energy back to EL2 is possible. The barrier between the two configurations is $\sim$0.34 eV. What makes EL2 important is its ability to fully compensate moderate concentrations of acceptors in GaAs and to pin the Fermi level near the middle of the band gap, yielding the material semi-insulating. Such crystals can be used as nonconducting templates for very high-speed GaAs circuitry.

Over the past two decades other metastable defects have been discovered, too many to describe. Up-to-date accounts of these multifaceted defects can be found in the proceedings of the biannual International Conferences of Defects in Semiconductors (1).

Outlook

Defects in semiconductors will continue to attract scientific and technical interest. New semiconductors, new hetero-interfaces, and new device structures are the sources of defects with a variety of electrical, optical, and structural properties. The properties will have to be understood so that the defects can either be eliminated or put to good use. The shrinking dimensions of semiconductor structures will require ever more sensitive and selective characterization tools. Has defect research reached its plateau? Hardly. Just watch the bustling activities surrounding strained layer structures, nanocrystals, or the group III nitrides and their alloys, to name a few.

References and Notes

1. For recent summaries, see Mat. Sci. Forum 258-63 (1997); Proceedings of the Eighth International Conference on Shallow-Level Centers in Semiconductors, Montpellier, France, 27 to 30 July 1998 [Phys. Status Solidi (B), in press].
2. H. J. Queisser, Mat. Res. Soc. Symp. Proc. in press.
3. C. Wetzel et al., Phys. Rev. Lett. 78, 3923 (1997).
4. A. K. Ramdas and S. Rodriguez, Rep. Progr. Phys. 44, 1297 (1981).
5. S. T. Pantelides, Ed., Deep Centers in Semiconductors: A State-of-the-Art Approach (Gordon & Breach, Lausanne, Switzerland, ed. 2, 1992); A. M. Stoneham, Theory of Defects in Solids (Clarendon, Oxford, 1975).
6. W. Shockley, IEEE Trans. Electron. Devices ED-23, 611 (1976).
7. S. Nakamura, Science 281, 956 (1998).
8. J. Neugebauer and C. Van de Walle, Phys. Rev. B 50, 8067 (1994).
9. H. R. Huff et al., Jpn. J. Appl. Phys. (Part 1) 37, 1210 (1998).
10. J. P. Hirth and J. Lothe, Theory of Dislocations (Wiley, New York, 1982).
11. H. J. Queisser, Mat. Res. Soc. Symp. Proc. 14, 323 (1983).
12. S. Nakamura and G. Fasol, The Blue Laser Diode (Springer, Heidelberg, 1997).
13. A. R. Smith, R. M. Feenstra, D. W. Greve, J. Neugebauer, J. E. Northrup, Phys. Rev. Lett. 79, 3934 (1997).
14. K. Graff, Metal Impurities in Silicon (Springer, Heidelberg, 1995).
15. G. Obermeier, J. Hage, D. Huber, J. Appl. Phys. 82, 595 (1997).
16. For a recent survey, see D. Gilles and A. Ohsawa, in Semiconductor Silicon 1998, H. R. Huff, U. Gösele, H. Tsuya, Eds. (Electrochemical Society, Pennington, NJ, 1998), vol. 98-1, p. 1047.
17. E. E. Haller and F. S. Goulding, Handbook on Semiconductors, C. Hilsum, Ed. (Elsevier, New York, ed. 2, 1993), vol. 4, pp. 937–963.
18. E. E. Haller, Festkörperprobleme: Advances in Solid State Physics XXVI, P. Grosse, Ed. (Vieweg, Braunschweig, Germany, 1986), pp. 203–229.
19. J. T. Walton, R. H. Pehl, Y. K. Wong, C. P. Cork, IEEE Trans. Nucl. Sci. 31 (no. 1), 331 (1984).
20. E. E. Haller, Phys. Rev. Lett. 40, 584 (1978).
21. ———, G. S. Hubbard, W. L. Hansen, IEEE Trans. Nucl. Sci. 24 (no. 1), 48 (1977).
22. See articles in J. Pankove and N. Johnson, Eds., Hydrogen in Semiconductors: Semiconductors and Semimetals (Academic Press, Orlando, FL, 1991), vol. 34.
23. S. J. Uftring, M. Stavola, P. M. Williams, G. D. Watkins, Phys. Rev. B 51, 9612 (1995); J.-U. Sachse, E. O. Sveinbjornsson, W. Jost, J. Weber, H. Lemke, ibid. 55, 16176 (1997).
24. I. Pomeranchuk, J. Phys. (Moscow) 6, 237 (1942).
25. T. H. Geballe and G. Hull, Phys. Rev. 110, 773 (1958).
26. E. E. Haller, Appl. Phys. Rev. in J. Appl. Phys. 77, 2857 (1995).
27. A. J. Mayur et al., Phys. Rev. B 49, 16293 (1994).
28. H. Bracht, E. E. Haller, R. Clark-Phelps, Phys. Rev. Lett. 81, 393 (1998).
29. P. A. Stolk, D. J. Eaglesham, H.-J. Gossmann, J. M. Poate, Appl. Phys. Lett. 66, 1370 (1995).
30. S. P. Murarka, Silicides for VLSI Applications (Academic Press, Orlando, FL, 1983).
31. W. Walukiewicz, Appl. Phys. Lett. 54, 2094 (1989).
32. For reviews, see P. M. Mooney, Semicond. Sci. Technol. 6, B1 (1991); Appl. Phys. Rev. in J. Appl. Phys. 67, R1 (1990).
33. D. V. Lang and R. A. Logan, Phys. Rev. Lett. 39, 635 (1977).
34. M. Mizuta et al., Jpn. J. Appl. Phys. 24, L143 (1985).
35. D. J. Chadi and K. J. Chang, Phys. Rev. Lett. 61, 873 (1988).
36. G. D. Watkins, Mat. Sci. Forum 38-41, 39 (1989).
37. J. Wolk et al., Phys. Rev. Lett. 66, 774 (1991).
38. J. Dabrowski and M. Scheffler, ibid. 60, 2183 (1988).
39. D. J. Chadi and K. J. Chang, ibid., p. 2187.
40. We gratefully acknowledge suggestions and help from A. R. Smith, R. M. Feenstra, and D. W. Greve of Carnegie Mellon University for providing us with Fig. 3 before publication; C. Van de Walle of Xerox, Palo Alto Research Center, for discussions and permission to reproduce Fig. 2; and J. Weber of the Max-Planck-Institut, Stuttgart, Germany, and M. Stavola of Lehigh University for reproducing Fig. 5. W. L. Hansen and W. Walukiewicz at Lawrence Berkeley National Laboratory (LBNL) have contributed enlightening discussions. H.J.Q. is grateful for the hospitality extended to him at the University of California at Berkeley and at LBNL.

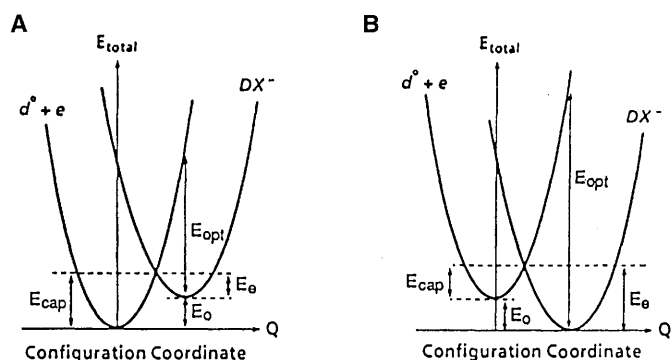**Fig. 7.** Configuration coordinate diagrams for DX centers in $Al_xGa_{1-x}As$; (A) $x < 0.22$, (B) $x > 0.22$.