Part I

Plenary Review

ORDER AND DISORDER IN SEMICONDUCTORS

HANS J. QUEISSER
Max-Planck-Institut für Festkörperforschung,Heisenbergst.1, D-70569 Stuttgart,Germany

ABSTRACT

The astounding success of microelectronics rests on a simple materials principle : creating a highly purified and perfected , spatially ordered semiconductor matrix , whose electrical and optical properties may be selectively adjusted by local substitutions of host atoms by dopant atoms. This unique materials utilization differs remarkably from all earlier technologies , because the controlled , almost imperceptibly small disorder by doping (rather than the ordered host) dominates the relevant properties ! Defect control is thus a major concern for semiconductor technology. Homogeneity is an absolute necessity for this strategy, but only a few of the semiconductors can be made so homogeneous as to suppress the strong deleterious effects of inhomogeneity. Recent advances are summarized : atomic resolution of defect analyses, multiatom reactions and hope for applications,contactless measurements, gettering as well as detailed theory of simulations. The emergence of novel quantum devices, with both reduced dimensions and reduced dimensionalities heralds a paradigm change, since the quantizing small geometries exert stronger influences than defects do; nevertheless, materials perfection and interface control remain prerequisites for these structures.

INTRODUCTION

The pervasive impetus of modern semiconductor technology has become an accepted fact. Scientific mastering of materials and processes has increased tremendously within a rather short time frame. Technological control has been derived from this scientific base, and an industry with more than $ 100 billions worth of silicon devices per annum- in 1994- and still almost incredibly high growth rates of production and applications is an economic reality as well as a matter of international industrial policy. The materials aspect of using perfected single crystals and applying local doping-control provide the basis of this unusual success. Earlier usage of materials differed remarkably. Bronze or steel are used for their specific bulk properties, the shaping and connecting of pieces is at the heart of iron-age or bronze-age technologies. Integration inside a regular spatial array of a host crystal is the semiconductor principle. The "Royal Road" to modern microelectronics consists in initially procuring a perfected single-crystal host, then locally establishing electrical and optical properties inside the host by specific replacements of host atoms by foreign "dopants". The somewhat disparaging expression of *defect* as a generic term for all deviations from the host perfection does not really convey the power of this "doping doctrine" for semiconducting materials.The early pioneers of germanium and silicon, however, placed great emphasis on the experimental verification that *n*-type and *p*-type doping by elements of the adjacent columns in the periodic table were accompanied by changes in the lattice parameter : atomic substitution of the host atoms as the guiding principle ![1,2]

This keynote speech shall attempt to offer a few recent examples of ongoing research to illustrate the power of defect utilization and the degree of sophistication that has been achieved .

Mat. Res. Soc. Symp. Proc. Vol. 378 © 1995 Materials Research Society

A warning is further uttered here, since this principle of doping by replacement can only be safely practiced for well-controlled , homogeneous materials - of which we still have only few semiconductors. Uncritical transfer of this principle to inhomogeneous materials does not guarantee success. Inhomogeneous materials derive many of their electronic properties *not* from the substituted point-defects but from the capricious details of the inhomogeneity itself. A similar situation arises in the current quests for novel quantum devices, where geometry plays an overriding role in establishing energy levels, densities of electronic states , or other electronic properties, thus diminishing the influence of substituents. Dimension-controlled quantum device structures , however, also command highly perfected base materials !

Dopant atoms, their interactions with native defects and their diffusion properties are exemplified first. Multiatomic defects, presently of vivid interest, and their stabilities are touched upon next. The topic of gettering, still of irreplaceable technical significance, is mentioned subsequently. The themes of inhomogeneity and quantum structures conclude this article.

CURRENT RESEARCH TOPICS

Dopants, Their Interactions and Diffusion

The diffusion of a simple trivalent acceptor or pentavalent donor in a silicon host still constitutes an essential element of device processing for bipolar and even for unipolar devices. An amazing degree of empirical knowledge rules these processes until today, because the atomistic details of diffusion -even for these simplest substitute atoms - are in reality very complex. A multitude of dopant-interactions with native defects, especially vacancies and interstitialcies ,govern the diffusion processes and thus determine the technologically so vital properties,such as depth of a junction, space charge widths, or channel properties. Diffusion of the donor phosphorous near an interface SiO_2 / Si provides a good example. The solubility of P in the oxide is lower than in silicon itself, therefore one expects a pile-up of P atoms in the Si near an interface to the oxide, well known from the pioneering work by Grove [3]. Yet, no really satisfying model simulation with reliable, predictive power was available for a long time. Lau and his colleagues [4] at the Siemens Laboratories then proposed a new model, which included a third phase, sandwiched in between oxide and elemental silicon. This hypothetical interface phase was assumed to provide ample amounts of sites for phosphorous atoms; such P atoms are then assumed to become electrically inactive, which is an easily acceptable proposition since the direct substitution inside a regular lattice no longer prevails. Lau *et al.*then proceeded to simulate diffusion conditions with a set of rate equations and found good agreement with experimental data. One single curve was presented to fit data for different P concentrations, a variety of temperatures and diffusion ambients, including inert anneals.[4] Diffusion process technology was thus satisfactorily explained with the appropriate parameters inserted, all based upon the *ad hoc* assumption of an intermediate phase.

Experimental verification of the postulated intermediate region was difficult because of the different forms of incorporated P. A chemical analysis , such as by secondary-ion-mass-spectroscopy, gives the total amount of phosphorous atoms. The electrically active, substitutional atoms, however, are the ones determining device characteristics; their amount may be almost one order of magnitude lower than the total. Nicollian and Chatterjee [5] have just recently provided such proof, thus explained the almost paradoxically appearing observation of

an *increasing* P concentration from the interface toward the silicon interior. Figure 1 shows the results obtained in the fit with the model; the profile parameters are adjusted toward best fit,also a diffusion length (0.95μm) is a result of such fit. This recent example of work pertaining to a seemingly very simple phenomenon, on which much of modern device technology has to rely, demonstrates first the inherent complexity and secondly the methodology of high-resolution characterization combined with detailed mathematical simulation in a multi-parameter space.

Impurities at Internal Interfaces

Evidence for such a novel intermediate phase with special properties has very recently [6] been obtained in a rather different experiment. Grain-boundaries were investigated in silicon crystals with the oxygen content typical for crucible - grown Si material,as is utilized in modern MOS-technology. The precipitation of this deliberately introduced foreign atom is a very important part of the gettering technology to remove unwanted impurity atoms. Very-high-resolution transmission electron microscopy revealed beautiful examples of amorphous Si-O phases,nucleated directly at the grain boundary. The border between precipitate and host matrix was always found to be a well-defined crystallographic (111)-type plane. However, Gatts *et al.*[6] attempted an analysis of yet higher resolution by means of neural - pattern -recognition methods for the obtained electron - energy - loss spectra.The result of this novel type of nonsubjective analysis by neural-network-technique was an existence proof for a new phase of a different chemical environment , again as sandwiched between crystalline Si and amorphous oxide.

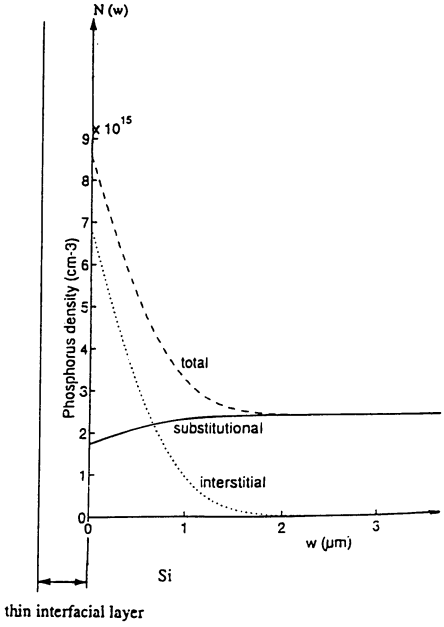

Fig.1 Profiles of various species of P *dopants in silicon near an oxide interface,see Refs.[4,5]*

This recent example of modern transmission electron microscopy in conjunction with refined mathematical analysis and simulation furnishes impressive proof for the power of present - day tools and the degree of sophistication: resolution and image-handling!

Another piece of evidence arose from Ref.[6], which corroborated very old results of my own [7], regarding grain boundaries, their structure and electrical properties. A naive model views a

grain boundary as a regular array of individual dislocations, which in turn supply electrically active sites -due to dangling bonds- along the dislocation line. A grain boundary (and even one single dislocation) should therefore be a strong electronic perturbation as a recombination,possibly even a doping ,site. This simple picture is, fortunately for device technology, incorrect. It is much more the combined effect of boundary plus its impurity atmosphere ,which affects the electronic properties. Measurements of the photovoltage arising at grain boundaries showed that the sign and the magnitude of such photo-signals could be massively altered by heat treatments,which are known to cause different phases of the oxygen impurities in Si. This old proposition of 1962 [7] was a strong deviation from the simple reductionist tenet of ascribing electronic behavior exclusively to point-like defects in the semiconductor lattice. Dislocations are defects contributing intensely to formation of many types of inhomogeneities. Remarkable efforts were exerted very early in semiconductor history to completely eliminate dislocations-for rather different reasons-in silicon, thus esta - blishing a very homogeneous base material of ever increasing wafer diame - ter,soon to reach 300 mm !

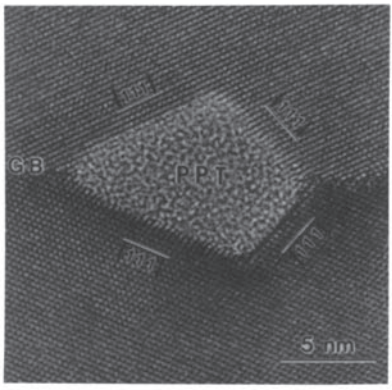

Fig.2 Electron micrograph of grain boundary in Si with oxide precipitate.Ref.[6]

Dopants and Native Defects

The elementary step for a diffusing impurity or dopant is the jump to a new site,which may be a substitutional or an interstitial site. Availability of vacancies or of interstitials therefore quite strongly affect jump probabilities. Any local nonequilibrium of these native defects thus severely affects the diffusion constant.This dependence plus the varieties of diffusion mechanisms greatly complicated the quantitative understanding of diffusion,even until today. Earliest evidence, for example in the growth kinetics of oxidation-induced stacking faults [8,9] or the very puzzling and disturbing phenomenon of the *emitter push effect*, harmful for small base widths in bipolar devices, indicated already that changes in the native defect densities had very serious consequences. This situation has been greatly improved during the last decade; we now have much better insights . For example, at least a semi-quantitative knowledge exists about the excess of interstitials being generated by oxidation of a surface, as compared to a nitridation. Technologies are being developed to not merely control dopant quantities, but also to get native defects under control. [10] . Such control is -once again-most urgently needed around the gates and channels of MOS silicon devices. Control of geometry and dopant is a strict necessity. A judicious choice of mixing of oxides and nitrides has been shown[10] to offer just the desired amounts of native defects to achieve specific channel conditions. Such advances indicate the sophistication reached in coupling control of vacancies and interstitials with those of the dopants themselves. A stacking fault grows and shrinks with fluxes of vacancies and interstitials , therefore one can actually utilize[11] such stacking faults as (almost quantitative) detectors for native defects ! Lateral and vertical distributions of Si interstititals have been monitored in this fashion in wafer-bonded samples; the interstitials were injected by dry oxidation.

Another example involves the non-doping group IV-element carbon and its influence on the diffusion of the common substitutional acceptor boron in silicon. Stolk and colleagues [12] incorporated C into B-doped Si and implanted Si excess self-interstitials by ion beam techniques.Carbon was shown to completely suppress the interstitialcy-enhanced boron diffusion; the substitutional carbon hence traps and deactivates the interstititals in crystalline Si !The diffusion processes thus become more scientifically scrutinized, which is an absolute necessity for today's submicron device design lengths.Dopant impurity diffusion processes must be meticulously controlled, which includesguiding of the native host-defects, too.

Limits of the Doping Doctrine

The elemental semiconductors and their substitutional dopants, such as B or P, represent an unusually favorable case. Within wide ranges, both *n*-type and *p*-type doping are here possible In almost all other cases for host and guest atoms, the situation is not nearly as auspicious.Rare earth atoms, for example, are very difficult to incorporate into silicon or gallium arsenide[13].Their valence mismatch reduces their solubility ; electrical pumping of the efficient luminescence is not easily feasible; the active f-shell is well screened by outer electrons. A marriage of the rare-earth doped luminescing oxides with the convenient semiconductors remains a predicament.

A similar difficulty arises for the more ionic semiconductors of the type A^{II} - B^{VI} ,whose larger bandgaps and stronger tendencies for radiative recombination are most attractive for optoelectronic applications. Doping is , however, exceedingly more difficult than for the elements. This deplorable situation is still not understood, even less so is it under technological control. [14] There is possibly not one single, dominant physical or chemical reason for this dopability problem. Self-compensation is one assumption; the crystal lowers its total energy under a doping attempt by creating native defects, which are much more strongly charged in these ionic compounds than in Si or Ge.This self-compensation brings the Fermi level back to a position closer to the middle of the forbidden gap. This plausible principle does not, however, apply generally.In other examples, it seems to be just the low solubility which makes doping almost impossible.[15]. The remarkable progress in lasers based on Zn-containing A^{II} - B^{VI} materials is founded on rather clever ways of circumventing these basic problems.[16] Not surprising is the prediction that our central theme "order-disorder in semiconductors" will be attacked very strongly for these compounds in the near future.It will, however, be interesting to see how far the analogs and patterns of the elemental semiconductors will carry us. There exist strong differences: high-dislocation-densities, for example,are quite common in light-emitting diodes, based on gallium nitride.[17] Densities exceeding 10^{20} cm^{-2} were found to be common in GaN, yet all minority-carrier properties were much less affected than in Si ,GaP or related materials ! In this compound GaN, therefore, nonradiative recombination is much less enhanced by dislocations, possibly because of the more ionic nature of the bonding [17] or , maybe, because of different tendencies of precipitation around dislocation cores ?

Compensation by inadvertent presence of the opposite type of dopant is a serious and often conveniently ignored problem.[18]. For most technical applications, however, compensation is not a serious problem,although it represents a degradation of crystalline perfection by incorporating too many impurity atoms per desired function. Minority carriers will therefore suffer a reduction of their lifetimes. The most noticeable effect of compensating impurities, say donors in a *p*-type crystal, is the rather sharp reduction of the mobility μ at low enough temperatures,

where scattering at ionized impurities [19] is the dominating mechanism for the mobility. In Si, this regime is almost impossible to reach for device conditions,whereas for GaAs and related compounds, compensation is more influential. All these considerations show the unusually favorable physics and chemistry for silicon with its precious facility for doping. The integrity of the oxide SiO_2 and the dopabilities make Si *the* preferred semiconductor material.

Multiatomic Centers

A semiconductor crystal represents a solvent, in which multitudes of chemical reactions may occur, very similar to the well-known paradigms in aqueous solution. Defects with opposing charges, for example, may associate and dissociate akin to a salt in water. One of the best known examples for Si are the pairs of iron atoms with all the acceptors, such as FeGa. The reaction kinetics can be determined, energies of dissociation are measurable [20], just as in the familiar textbook knowledge concerning charged ions dissolved in water.

By far the most intensely studied biatomic centers in Si were recently those containing hydrogen as one of the partners.[21] Defects can easily be neutralized, thus rendered passive by the admixing of hydrogen. One recent example for these investigations is the work by Cheng and Stavola [22].They studied the reorientation of BH and BD complexes and found evidence for a thermally assisted tunneling. Figure 3 gives a result. The jump rates are plotted for the tunneling between two different bond-centered sites , as indicated by the arrow in the inset. These centers were initially oriented by the application of mechanical stress and then allowed to relax. Tunneling was inferred from the non-Arrhenius behavior. Notice from the figure that the deuterated sample is the one with quicker rates, at least for the tempera - tures shown ! Bonding and release make diffusion of H in Si and in other semiconductors a difficult problem and a highly complicated affair. Hydrogen is, however, ubi - quitous,not only in wet chemical processes. [42]

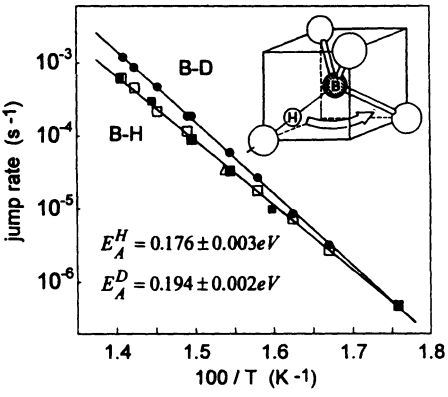

Fig.3 Tunnel rates of H and D in pairs with boron, from Ref. [22]

Multiatomic centers are in many cases multistable defects, such as indicated by Fig.3.They may assume different positions, for example with respect to hydrogen relative to the boron atom. One could imagine bistable centers, where the two possible configurations are more distinctly different from each other as in the BH case, for example via a difference in the inter - atomic distance. Such centers with strongly differing arrangements offer some hope for a utilization in memory storage. The two configurations might represent a one and a zero,they might be switched by externally applied signals and might be susceptible to read-out.Such an application would yield a very small and -hopefully- simple means of data storage in Si. However, the difficulties seem formidable.At present, observable bistability is limited to low temperatures, and the addressing seems problematical for optical access, let alone for an electrical input and

output. This topic of atomic-size data storage is, however, such a fascinating proposition that more effort and research seem probable for the future.

The Indispensable Art of Gettering

Vacuum tubes had to rely on *gettering* the residual gases by means of a pill of a reactive metal to bind oxygen and nitrogen. This practical act out of desperation has been carried over to standard silicon device technology. Inadvertent and unavoidable metallic impurities must somehow be eliminated from the (relatively small) regions of active device relevance and ought to be firmly fixed in some passive region. No *hard p-n* junction , i.e. one with a sharp reverse I-V curve, was feasible, at least in the earlier history of bipolar devices, without such a gettering step, which removed the metallic generation centers from the junction's space charge layer.[23] Metallic impurities , often from the furnaces or implantation equipment, are still very difficult - and expensive - to remove by any other technique than by some gettering. The sinks for these metal atoms can be manifold: glassy surface layers, dislocations or even more robust, cruelly introduced lattice damage on the wafer-backside,[24] and -most importantly- the convenient oxide precipitates. Such gettering seems to apply quite universally for most metallic impurities [25] . The local change in the metal-solubility must be sufficient to capture the yield-diminishing culprit, and the diffusivity of the metal must suffice in order to reach the sink. These conditions can usually be met quite easily, and the needed process step can in most cases be inserted into the device processing sequence without great problems. Still, the details of the gettering kinetics and the interactions, especially with the native defects, are quite complex.Gettering is in principle a reversible process[26] ;care must be taken to avoid re-dissolution. The smaller dimensions of future Si devices will here require more of a scientific basis; supportive research will be needed. Such research will, however, have to be based on a very close interaction between researchers in fundamental materials science and device production-line practitioners.

Pernicious Inhomogeneities

The powerful dogma of point-defects inducing semiconductor properties relies on a uniform, homogeneous crystal base. As soon as inhomogeneity of any sort creeps into the matrix, our principles begin to fail. Compensation by minority dopants or the deleterious, ill-defined influence of dislocations or two-dimensional defects ,such as stacking faults or boundaries,even surfaces, have already been mentioned here. The basic theory of inhomogeneous materials and a discussion on their foundations upon quantum theory and thermodynamics have been outlined by Pantelides [27]. Detailed mathematical treatment of transport and other electronic properties in inhomogeneous semiconductors has been given by Pistoulet and his students [28]. The situation is demonstrated to quickly become intractable; general conclusions are very difficult to be formulated. Inhomogeneities have one major deleterious consequence :local potential fluctuations haphazardly arise ! Transport is therefore randomly affected, which is a serious competition against the willful, designed external potentials, such as on a field-effect gate, to control electronic current flow. The function of a device, especially one of small geometric dimensions, is lost. The concept of doping as the source of electronic behavior breaks down .

One particularly forceful example of this breakdown is established by the maverick phenomenon of *persistent conductivity*. Many semiconductor samples of questionable homogeneity and ill-defined doping display a persisting conductivity *after* a photoconductivity-inducing illumina-

tion has been switched off.[29] Such a sample appears to have some sort of barrier against recombination, thus artificially prolonged lifetimes [30]. Such effects have long been known and have been associated -in most cases- with a lack of homogeneity.[29] For example, in a semiconductor with clusters of radiation-induced defects, space charges arise,as indicated schematically in Fig.4.[28-31].Upon illumination, one type of carriers is trapped at these radiation-induced defects, the other type is mobile to conduct current, which prevails even after illumination ceases. The cluster thus separates localized from delocalized carriers,very much like a gate.Any potential fluctuation fulfills this function, just like a junction does in a solar cell, separating electrons from holes.Persistence had therefore initially been taken as evidence for all kinds of non-uniformities [31].

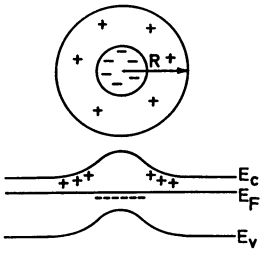

Fig.4. Schematic of a carrier-separating fluctuation ,Refs.[30,31]

This generally accepted, but not much recognized viewpoint underwent a complete change with the detection of specific multiatomic centers in the compound semiconductors, especially gallium arsenide. The so-called DX-center, a donor with a then unknown second partner was interpreted [32] as being bistable and being the source of the excess electrons persisting after illumination. The doping-doctrine was here successful once again. Atomic-size defects were linked to an unusual conduction property: the liberated electrons are unable to recombine with their source defects since they cannot overcome a configurational barrier at low temperatures. Individual defects in an orderly semiconductor had now found another significant property, aside from doping or lifetime-dominating recombination. This success in GaAs was swiftly carried over to many other materials, which were of much lesser homogeneity than GaAs. Suddenly, any observation of persistent photoconductivity was ennobled as an indicator of a most interesting defect,usually associated with strong lattice-interactions, rather than being an embarrassing stigma of inhomogeneity[30]Yet, the fact remains that any type of potential fluctuation will generate persistence under quite general experimental conditions. Our own work on GaAs epitaxial layers of high perfection has demonstrated this principle of spatial carrier separation [33-35] , with one type of carriers trapped, the other being free to carry a current. The accumulation of carriers is quantitatively explained with this idea [34]. A specially made GaAs sample with two differing trap densities demonstrates two different slopes in a curve of current *versus* photon dose , which is a rather convincing proof [30,35].see Fig.5 We always warn against too easily interpreting such persistence phenomena with some multistable, special defect.The first,rather prosaic and admittedly discomfiting duty has to be a thorough check against potential fluctuations.Let me reiterate the message that the doping doctrine must be based on homogeneity !

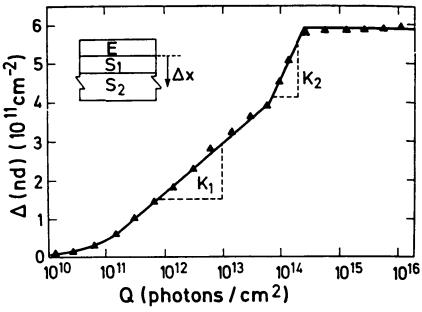

Fig.5 Dose-dependence of persistence in a two-layer trap configuration Ref.[35]