PHYSICAL REVIEW B VOLUME 22, NUMBER 7 1 OCTOBER 1980

Complex nature of gold-related deep levels in silicon

D. V. Lang
Bell Laboratories, Murray Hill, New Jersey 07974

H. G. Grimmeiss and E. Meijer
Department of Solid State Physics, Lund Institute of Technology, Box 725, S-220 07 Lund 7, Sweden

M. Jaros
Department of Theoretical Physics, The University, Newcastle upon Tyne, England
(Received 12 October 1979)

We report detailed measurements of the thermal and optical properties of the well-known gold-acceptor level in silicon. From a comparison of these properties measured in two types of silicon p^+n diodes (one fabricated from epitaxially grown silicon and the other from silicon grown by the Czochrolski technique), we find that there are at least two (and perhaps more) different types of gold-acceptor centers in silicon. Measurements of the ratio of the deep-level transient spectroscopy signal due to the gold-donor level to that of the gold acceptor in the same sample show that these two levels are not related to the same gold center, as had previously been believed. A comparison with data in the literature for five other approximately midgap levels in silicon (Ag, Co, Rh, S, and process-induced levels) shows that the electron thermal-emission rates of these are all identical to that of the gold acceptor within experimental error. This suggests that these deep levels have the same underlying defect structure. Comparisons of our measured electron-capture cross sections with those reported in the literature for the gold acceptor show a heretofore unreported correlation between the magnitude of this cross section and the ratio of gold concentration to that of the shallow donor impurity. This suggests that ion pairing between gold acceptors and shallow donors plays a role in determining the capture cross section of the defect. We also have measured for the first time the capture cross sections associated with the gold-donor defect in n-type material and find results significantly different from those reported in the literature for p-type silicon. The optical cross sections of the gold acceptor show nearly a factor of 10 difference in magnitude and subtle differences in shape between epitaxial and Czochrolski samples. It is therefore possible that oxygen is also playing a role in the gold-based defect complexes. Attempts to obtain the temperature dependence of the gold-acceptor level from thermal capture and emission data proved inconclusive and led to various paradoxes which we discuss. Finally, we discuss possible models for gold-related defect complexes which are consistent with our results and might form the basis for future work.

I. INTRODUCTION

The deep donor and acceptor levels in gold-doped silicon have been perhaps the most widely studied of any deep level in semiconductors. Refinements of experimental technique have led to numerous studies claiming increased accuracy. Yet the values quoted for energy levels and capture cross sections in the literature seem to show larger variations than should be expected from the quoted limits of error. For example, as Engstrom and Grimmeiss[1] have pointed out, the quoted values for the depth of the acceptor level from the conduction band vary from 0.49 to 0.58 eV. Similarly, Kassing and Lenz[2] have shown that the reported values of the electron-capture cross section for the acceptor fall within a range of nearly a factor of 50 between the largest and smallest values.

Attempts to determine the temperature dependence of the acceptor energy level have also lead to controversy. Engstrom and Grimmeiss[1] carefully analyzed the temperature dependence of the thermal capture and emission of holes at the acceptor and came to the conclusion that this transition has the full temperature shift of the entire band gap. It was therefore argued that the acceptor level was fixed to the conduction band as a function of temperature, in agreement with the observation[3] that the thermal shift of the optical-emission line shape for electrons was much smaller than that for holes. Recently, however, Brotherton and Bicknell[4] analyzed the temperature dependence of the rates of thermal capture and emission of electrons at the acceptor level and concluded that this transition has the full temperature shift of the gap. Therefore, they argued, the level must be fixed to the valence band.

In an effort to find a mechanism for the large, temperature-independent electron-capture cross section at the neutral acceptor level, Lang and Jaros[5] analyzed the electron optical-emission line shape of this level searching for evidence of thermal broadening due to the electron-phonon coupling. Such broadening would have supported the multiphonon emission (MPE) mechanism of capture.[6,7] The resulting data[5] indicated considerably more thermal broadening of this transition than had previously been reported.[3]

 © 1980 The American Physical Society

The rather confused situation we have just outlined for data on the acceptor level in gold-doped silicon can be explained by assuming either

(1) that there are significant, unrecognized systematic errors in some of the experimental techniques which have been employed, or

(2) that the "gold acceptor" is not a well-defined unique level but rather a family of closely related but different gold-based complexes.

This latter possibility is suggested by the recent results concerning Cu-related complexes in GaP.[8] In order to determine the extent to which these two possibilities might be playing a role in the various experimental discrepancies in gold-doped silicon we undertook a joint project to study various types of gold-doped samples using a wide variety of different junction space-charge techniques. Our major emphasis was on a study of the acceptor level. We found that the various techniques gave essentially identical results, namely, that the different samples did indeed have subtle but significant differences in the thermal and optical properties of the acceptor level. Thus there is not a single, well-defined gold-acceptor defect in silicon, as had been universally assumed, but rather a family of gold-related defects with slightly varying properties.

The evidence we will present which supports the view that gold in silicon is a family of complex defects is detailed and perhaps somewhat confusing. In particular, we make use of our new data as well as results taken from the literature to build our case. Therefore we will discuss both new and old data throughout the paper whenever such data are relevant to the subject at hand. To assist the reader in following the detailed arguments and the presentation of the data we will outline here the main arguments and conclusions.

First we find from a careful analysis of deep-level transient spectroscopy (DLTS) spectra that the gold-donor and gold-acceptor levels are not of equal concentration in all samples. This implies that the donor and acceptor are not related to the *same* gold center, as had been universally assumed in the past. Thus freed from the constraint of dealing with the donor and acceptor levels together, we focus the remainder of the paper on the gold acceptor as a separate class of defect complexes. In comparing the electron thermal-emission rate for the gold acceptor with data in the literature for other deep levels in silicon, we find that there is a group of six defect systems which have thermal-emission properties indistinguishable from the gold acceptor. Since the literature results typically were carried out with great care using undoped control samples, it is probably safe to rule out extraneous gold contamination. Thus we propose that these six systems (Au, Co, Ag, Rh, S, quenched defects) may all have a defect of the same basic structure and properties but associated with a different chemical impurity in each case.

We contrast this striking similarity in electron thermal-emission rates for *different* impurity systems with small but significant differences in other properties for the *same* impurity system (Si : Au). The two properties which show variations among different samples of gold-doped silicon are the electron–thermal-capture cross section and the optical cross sections for photoionization of electrons or holes from the gold acceptor. In contrast to the thermal-emission rate, which is primarily a property of the ground-state energy of the system, the capture and optical cross sections are much more strongly influenced by the shape of the defect wave function and possible excited states. By adding our new capture data to a compilation of previous measurements taken from the literature we note a correlation which has never before been reported, namely, that the electron-capture cross section of the gold acceptor varies by nearly a factor of 50 depending on the ratio of the gold concentration to that of the shallow donor. This suggests that the gold complex may also be changed in subtle ways by the association of phosphorous. Finally, the subtle differences we see in the optical cross sections of gold centers in epitaxial versus Czochrolski samples suggests that some additional impurity such as oxygen may also be a part of the complex in some samples.

The picture we hope to leave is one of a complex defect structure for the gold acceptor, but not hopelessly complex. The overriding similarities between Au, Ag, Co, Rh, S, and quenched-in defects far outweigh the much more subtle effects due to the associations with phosphrous and (perhaps) oxygen. Indeed, such broad similarities among classes of defect complexes and the dominant role of the lattice in determining the properties of these complexes is an emerging feature of our current theoretical understanding of deep levels in semiconductors.

In Sec. II we will describe the different samples used in this joint project and briefly outline the experimental techniques. The thermal emission and capture rates of electrons at the acceptor will be discussed in Secs. III and IV, respectively. In Sec. V we will present the optical cross-section data on our two major types of samples. An analysis of the thermodynamics of energy levels will be presented in Sec. VI with a special emphasis on the importance of the often overlooked entropy factors. Section VII is a discussion of the conclusions which may be drawn from the data, including some simple

speculative models which may be helpful in the search for the underlying structural difference between the various types of gold defects. In Sec. VIII we summarize the major points of the paper. Finally, we place in the Appendix an analysis of electron and hole capture at the gold-donor level which is necessary to support the validity of concentration measurements (in Sec. III) which show that the donor and acceptor levels are not related to the same defect.

II. EXPERIMENTAL DETAILS

Two different types of gold-doped silicon p^+n junctions were used in this study. The first, which we will call type A, was prepared from epitaxial silicon and is the same type as was used by the Lund group in previous optical[3,9] and thermal[1] measurements. The n-type layer was a lightly doped 10-μm-thick epitaxial layer $[N_D - N_A = (4-7)\times 10^{15}\ \mathrm{cm}^{-3}]$ grown on a p^+ substrate $[N_A - N_D = (5-7)\times 10^{18}\ \mathrm{cm}^{-3}]$. In these samples the doping profile near the junction is not perfectly abrupt but is slightly graded.[10] Various wafers were gold diffused at temperatures between 890 and 960℃ to give samples with a homogeneous gold concentration of from 5×10^{14} to $3\times 10^{15}\ \mathrm{cm}^{-3}$.

The second type of sample, which we will call type B, is an abrupt p^+n step junction formed by boron diffusion into 5 ohm cm phosphorous-doped silicon grown by the Czochralski technique. Gold was evaporated on the back of the wafer after the junction was formed and diffused uniformly through the sample to a concentration of $1\times 10^{14}\ \mathrm{cm}^{-3}$. The net donor concentration in a completed p^+n junction was found to be $8\times 10^{14}\ \mathrm{cm}^{-3}$. The Ohmic contact on the top of each type of sample was much smaller than the junction area to allow direct illumination of the space charge layer in the optical experiments.

The experimental techniques used in this work were various forms of junction space-charge spectroscopy. Both steady-state and transient methods were employed. Since we found that all methods gave essentially identical results when applied to the same sample, and since all of the methods have been discussed extensively in the literature, we will very briefly mention here the various techniques which were used.

The junction *current* methods have been reviewed by Grimmeiss.[11] One of the methods which we used to measure optical cross sections was the dual-light-source steady-state photocurrent method.[11,9] Junction *capacitance* methods have been reviewed by Miller *et al.*[12] for the thermal-emission case. We used fixed-temperature capacitance transients to measure precisely thermal-emission rates and deep-level transient spectroscopy (DLTS) to study the relative concentrations of donors and acceptors. The photocapacitance methods which we used have been recently discussed by Henry and Lang[7] and by Kukimoto *et al.*[13] The capture cross sections were measured by the bias-pulse method developed by Henry *et al.*[14] In using this method we were careful to avoid the nonuniform capture rate in the edge region of the space-charge layer which is a problem at low bias voltages.[15]

For the optical and thermal transient measurements we used a precision temperature-controlled Dewar in which the sample was cooled by direct contact with an exchange gas which had a thermal stability better than 0.1 K. The temperature of the sample probe was monitored by a copper-constantan thermocouple which was calibrated with respect to the actual temperature of the sample by mounting a GaAs diode thermometer in the normal sample position. Such painstaking calibration procedures were needed to ensure the accuracy of the thermal-emission–rate measurements. We have found that without such calibrations it is possible for the sample temperature to be in error by several degrees, particularly in less sophisticated Dewar systems. Furthermore, the temperature error often may not be constant but may depend on the sample temperature. This effect is a particular problem with cold-finger-type Dewars and can give rise to a systematic error in the thermal-emission activation energy as determined from the slope of an Arrhenius plot.

III. THERMAL-EMISSION RATES

As we pointed out in the Introduction, there is a rather wide range in the reported values for the activation energy of thermal-electron emission from the gold-acceptor level to the conduction band. The extreme values range from 0.49 to 0.58 eV.[1] However, when these data are all plotted together as in Fig. 1, one sees that the various values of the electron thermal-emission rate e_n^t all fall very close to the same straight line corresponding to an activation energy of 550 ± 5 meV (with T^2 correction). The data in Fig. 1 are taken from Sah *et al.*[16] who reported an activation energy of 547 ± 2 meV, from Yau and Sah[17] who reported 582 meV from Parillo and Johnson[18] who reported 490 ± 20 meV, from Engstrom and Grimmeiss[1] who reported 553 meV, and from Brotherton and Bicknell[4] who reported 555 meV, all with the standard T^2 correction. These e_n^t values lie within a range of a factor of 1.5 at the same temperature or a range of 3.1 K at the same emission rate.

Our measurements of e_n^t in both A- and B-type samples fall in the midst of the data presented in

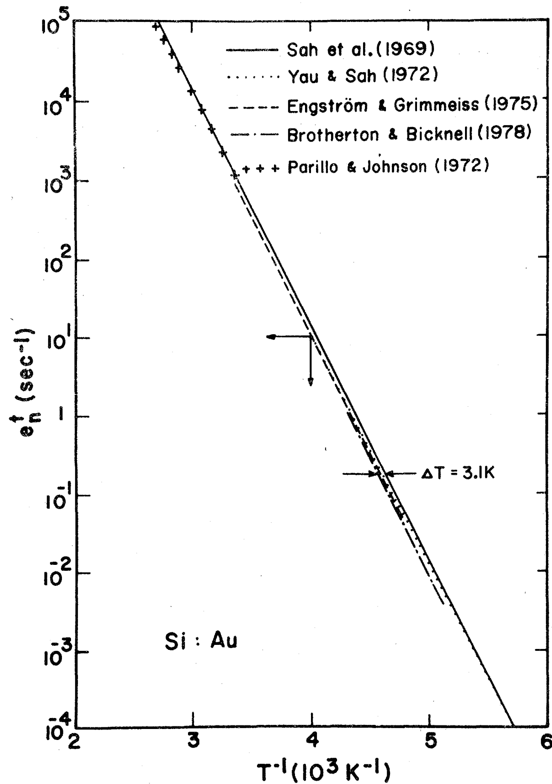

FIG. 1. Electron thermal-emission rate of the gold-acceptor defect versus inverse temperature for five studies reported in the literature.

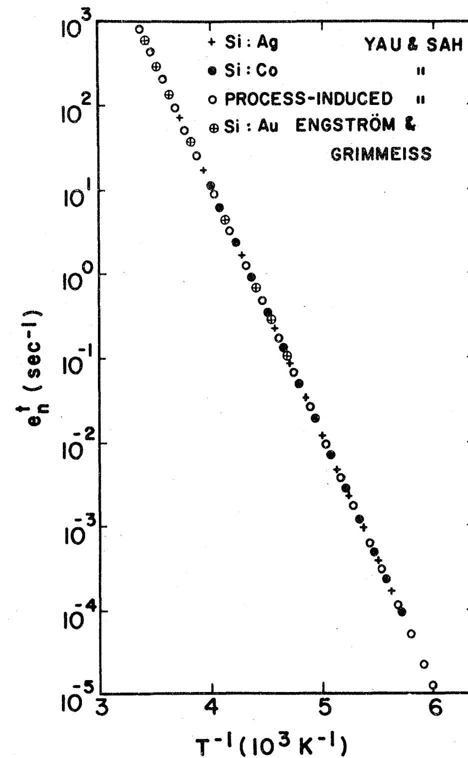

FIG. 2. Electron thermal-emission rates versus inverse temperature reported in the literature for the approximately midgap deep levels found in silicon doped with Ag, Au, and Co as well as the so-called process-induced defects found in samples quenched from 1100 °C. The various symbols indicate best fits to the Arrhenius plots in the literature and not the exact data points.

Fig. 1, almost exactly on those of Engstrom and Grimmeiss.[1] The activation energy which we obtain from these new data is 553 meV. Note that because of the well-known T^2 term in the emission rate, the slope of the raw data in Fig. 1 is actually $2kT_m = 43$ meV larger than the activation energy of 553 meV. This is the so-called T^2 correction with T_m chosen as the temperature of the midpoint of the temperature range of the data. The emission rate of sample B was only 4% higher than that of sample A at the same temperature. The effect of electric fields on e_n^t was also quite small, being less than 5% for a reverse bias change of from 2 to 20 V.

The close similarity between the thermal-emission rates of samples A and B is surprising in view of the significantly larger differences between these samples in their capture and optical properties, which we will discuss in later sections. Even more surprising, however, is the similarity in thermal-emission rates reported in the literature for several *different* impurities in silicon. This is strikingly shown in Fig. 2 where in addition to the gold-acceptor data[1] we have indicated best fits to the electron thermal-emission rates for the deep levels introduced near midgap in Ag-doped silicon,[17] Co-doped silicon,[17] and quenched silicon (process-induced levels).[19,20] Two other deep levels also have essentially identical e_n^t values when plotted on the scale of Fig. 2: the midgap levels in S-doped silicon[21] and in Rh-doped silicon.[22]

In spite of these e_n^t similarities, however, it is clear that these are distinguishably different defects. The most obvious difference is that the midgap levels introduced by the diffusion of Co, Rh, Ag, or Au into the sample are acceptors whereas the midgap levels in S-doped and in quenched silicon are donors. There are also significant differences in the electron-capture properties. For example, the electron-capture cross sections of the deep acceptors in Rh- and Au-doped diodes differ by a factor of 30.[22] Unfortunately, a detailed study of the various thermal and optical difference among these centers, such as we report here for the Au level in different samples, has not been done for the other five levels. Therefore we do not know the full extent of the differences between these defects.

The differences among samples doped with Co, Ag, or Au can most readily be determined by observing the thermal emission from the donor level

located in the lower half of the gap. The hole-emission activation energy (with T^2 correction) is 0.345 eV in Au-doped silicon,[16,22] 0.365 eV in Co-doped silicon,[17] and 0.393 eV in Ag-doped silicon.[17] In Rh-doped silicon there is no state in the lower half of the gap, but a second state (reported to be an acceptor) in the upper half of the gap with an electron-emission activation energy of 0.315 eV.[22] In order to show the thermal-emission signals from the donor and acceptor levels it is most convenient to use the DLTS technique.[12,23] Figure 3 shows DLTS spectra (50-sec^{-1} rate window) from our gold-doped samples of type A and type B. Also shown in this figure are the temperatures where the donor and acceptor levels should give peaks (at 50 sec^{-1}) in Co- and Ag-doped silicon. In an ideal p^+n diode electron (majority-carrier) emission gives a negative peak for a majority-carrier–pulse spectrum and hole (minority-carrier) emission gives a positive peak for an injection-pulse spectrum. However, since the A-type samples have slightly graded junctions, we also see a weak negative peak for the donor level on the p side of the junction, since hole emission is a majority-carrier process in p material. Such a negative peak is not seen in sample B, which is more nearly an abrupt

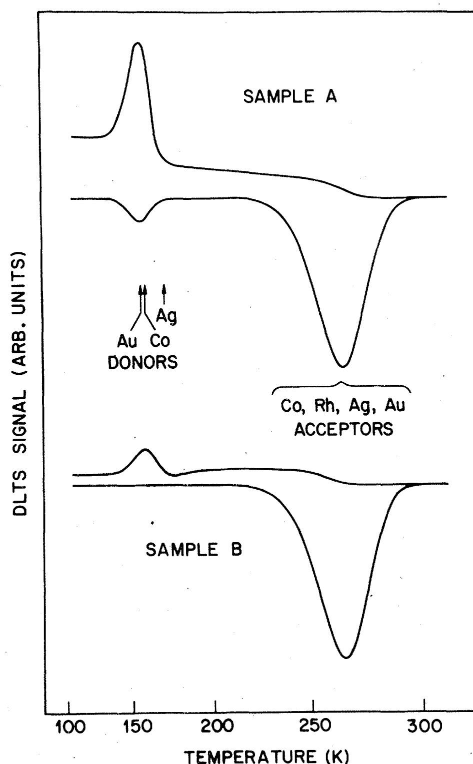

FIG. 3. DLTS spectra of our gold-doped p^+n diodes of type A and type B. Positive signals correspond to hole emission while negative signals correspond to electron emission. The rate window was 50 sec^{-1}.

p^+n junction.

Ever since the earliest Hall measurements on Au-doped silicon[24,25] it has been widely accepted that the donor and acceptor levels belong to the same substitutional gold impurity, i.e., gold has three charge states in the gap.[26] In the DLTS spectra of Fig. 3, however, it is quite clear that the donor level in sample B has a significantly smaller signal and hence lower concentration than the acceptor level. This implies that the donor and acceptor *cannot* be related to the same defect. Engstrom[27] has also recently seen unequal DLTS signals from the Au-related donor and acceptor levels. In this case the ratio of the donor to the acceptor signal was actually observed to vary during low temperature (~200 °C) annealing of the Au-doped samples. One must be very careful, however, in extracting concentration information from DLTS minority-carrier–emission spectra. It would be quite possible, for example, to fill only partially the donor level during an injection pulse and thus generate a DLTS signal which is *less* than the defect concentration. Since this paper is primarily concerned with the acceptor level, we give in the Appendix a detailed analysis of the trap-filling process at the donor level during an injection pulse and show that under the conditions of Fig. 3 the DLTS peak height is indeed proportional to the defect concentration, i.e., the donor is fully saturated with holes during the injection pulse. In the process of verifying the donor-to-acceptor concentration inequality we also find that the capture cross sections for the donors seen in our n-type samples are very different from those reported in the literature for donors in p-type material. Indeed, these p-type donors could be seen only by photocapacitance or photocurrent in n-type Si and not by DLTS. This further complicates the Au-doped Si picture and only serves to amplify the main point of this paper, namely, that Au in Si is not a simple defect.

In the past, the major evidence for gold being a three-charge-state system has been that the donor and acceptor states were of roughly the same concentration.[24,25] This is certainly a necessary condition but it is not sufficient to prove the existence of multiple-charge states. For example, Cu in GaAs has long been considered to be a double acceptor based on Hall-effect concentration measurements.[26] However, photoluminescence measurements of excitons bound to Cu in GaAs show that the two levels have different site symmetry and hence are most likely two different Cu-related defects.[28,29] Even more conclusive, however, are DLTS measurements which show the two Cu levels to have significantly different concentrations in the same sample.[30]

Another example of levels which might errone-

ously be assigned to a three-charge-state defect is the case of the two very common deep levels of unknown origin in GaAs which are usually referred to simply as level A and level B.[30] These levels almost always occur with very nearly the same concentration,[30,31] so that one might be tempted to think, as in Au-doped silicon, that they are two levels of the same defect. Recently, however, in several samples of GaAs these levels have been present in very different concentrations,[31] thus proving that they are not related. Apparently the occurrence of *different* defects with nearly the same concentration is not as unusual as one might expect. The common assertion that two deep levels of equal concentration most likely are related to the same defect is thus a very weak and potentially misleading argument. A good example of an experiment which actually proves the existence of a defect with three charge states is the combined photocapacitance and carrier-capture measurements on oxygen in GaP.[13,14]

IV. ELECTRON CAPTURE AT THE ACCEPTOR LEVEL

In contrast to the close similarities among e_n^t values for different gold-doped samples (and even for different impurities), the reported capture cross-section values deviate widely in the literature, lying in a range of over an order of magnitude for electron capture at the gold-acceptor level.[2] It is essential to have a direct measurement of this cross section in order to evaluate the thermodynamic variables associated with the electron-emission process. Therefore we measured the electron-capture rate at the gold-acceptor level in our two types of samples.

The electron-capture cross section σ_n is defined by

$$\tau^{-1} = \sigma_n v_{\text{th}} n + e_n^t , \qquad (1)$$

where τ is the time constant of the trap-filling process with a free-electron concentration n.[11,12] The average electron thermal velocity v_{th} is defined[26] as

$$v_{\text{th}} = (3kT/m^*)^{1/2} , \qquad (2)$$

where k is Boltzmann's constant, T is the absolute temperature, and m^* is the electron density-of-states effective mass. In our calculations of σ_n we have used[32]

$$m^* = (m_l^* m_t^{*2})^{1/3} = 0.35 m_0$$

so that in cm/sec

$$v_{\text{th}} = 1.97 \times 10^7 (T/300)^{1/2} . \qquad (3)$$

The capture properties of a deep level are also often expressed in terms of the capture coefficient c_n which is related to the capture cross section by

$$c_n = \sigma_n v_{\text{th}} . \qquad (4)$$

Our data for sample B are shown in Fig. 4. It is clear that c_n is proportional to the square root of temperature over a wide range (115–424 K). Therefore σ_n is independent of temperature, as has been found in all other reliable measurements (see Table I). The value of the cross section is $\sigma_n = 1.7 \times 10^{-16}$ cm^2. The deviation from a square-root temperature dependence in the capture rate at high temperatures is due to the e_n^t term in the measured time constant [see Eq. (1)]. The open circles in Fig. 4 correspond to the true value of c_n after e_n/n has been subtracted from τ^{-1}/n.

The electron-capture rate at the gold acceptor in sample A was measured at 295 K. The value of the cross section ($\sigma_n = 0.69 \times 10^{-16}$ cm^2) is 2.5 times smaller than that of sample B. This is outside the possible range of error in this type of bias pulse measurement. A very conservative estimate of the range of error in σ_n is certainly less than ±20%. This is, therefore, the first indication that the gold-acceptor levels in our two types of samples are actually different defects.

Our measured values for σ_n in samples A and B are compared with other measurements of the electron-capture cross section of the gold-acceptor level in Table I. In this table we have included only *direct* measurements of the capture rate. Thus we have specifically excluded values of σ_n derived from thermal-emission measurements by detailed balance arguments. As we will show in Sec. VI, detailed balance calculations are not valid unless the temperature dependence of the energy level is known. Previous workers have implicitly as-

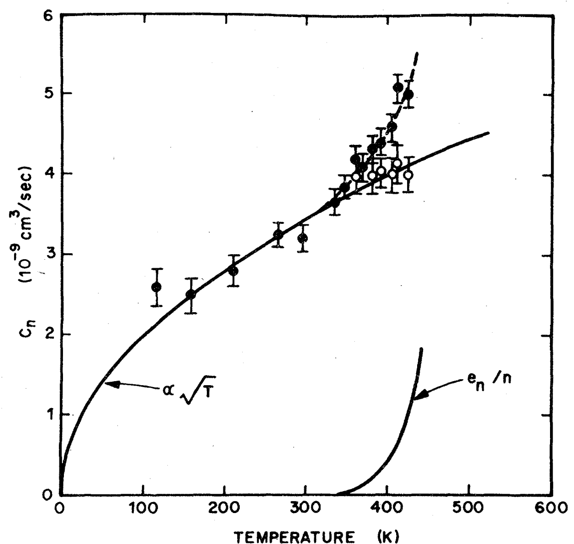

FIG. 4. Thermal-capture coefficient ($c_n = \sigma_n v_{\text{th}}$) versus temperature for electron capture at the gold acceptor.

TABLE I. Summary of electron-capture cross sections for the gold acceptor as reported in the literature.

Reference	Method	N_{Au}/N_D	Electron-capture cross section σ_n (10^{-16} cm^2)	Temperature range (K)
Fairfield and Gokhale (1965) (Ref. 34)	photoconductive decay	$\lesssim 0.1$	0.85	300
Pals (1974) (Ref. 22)	junction capacitance-pulsed bias	$\lesssim 0.17$	1.3	77–220
Barbolla *et al.* (1976) (Ref. 35)	junction capacitance-pulsed bias	0.18	0.83 [a]	80–205
Brotherton and Bicknell (1978) (Ref. 4)	junction capacitance-pulsed bias	<0.23	0.9	77–280
Sample *A*	junction capacitance-pulsed bias	<0.1	0.69	295
Sample *B*	junction capacitance-pulsed bias	0.1	1.7	115–424
Bemski (1958) (Ref. 25)	minority-carrier lifetime	0.45–0.88	5.0	200–450
Kassing and Lenz (1974) (Ref. 2)	SCLC	$\gtrsim 1$	4.7 [a]	300
Kassing *et al.* (1975) (Ref. 36)	SCLC	$\gtrsim 1$	7.3 [a]	172–316
Davis (1959) (Ref. 37)	majority-carrier lifetime	$\gg 1$	20	77
Nagasawa and Schulz (1975) (Ref. 38)	junction capacitance-bias pulse	$\gg 1$ (in ion-implanted region)	21.8	300

[a] These values differ from the original reference. They have been recalculated here with $m^*/m_0 = 0.35$ to be consistent with the rest of the table.

sumed without proof that the level was fixed to the conduction band[16,33,34] or else have attempted to incorporate a temperature dependence by arbitrarily assuming that the level had half of the temperature dependence of the gap.[18]

The data in Table I range over a factor of 30, from 0.69×10^{-16} to 21.8×10^{-16} cm^2. This large a deviation is far outside the expected range of errors in any of the measurements. In addition, there is no consistent variation with experimental technique. We believe, therefore, that these various values of σ_n are further evidence for different types of gold defects. If one assumes that all of the measurements are reliable, one sees that the data in Table I fall into three fairly distinct groups based on the concentration of gold N_{Au} relative to the shallow donor (most likely P) concentration N_D. The first group,[4,22,35,36] which includes our samples *A* and *B*, has $N_{Au} \ll N_D$, typically $N_{Au}/N_D \approx 0.1$. This group has the smallest cross-section values [$(0.69-1.7)\times10^{-16}$ cm^2]. Next, a group of intermediate cross sections [$(4.7-7.3)\times10^{-16}$ cm^2] corresponds to samples in which $N_{Au} \approx N_D$.[2,25,37] Finally, the largest reported cross sections [$(20-21.8)\times10^{-16}$ cm^2] have all been measured using samples in which $N_{Au} \gg N_D$.[38,39]

This correlation of cross-section values with doping situations where either the gold or the shallow donor is the dominant species suggests that some sort of ion-pairing phenomenon might be taking place. We will discuss this possibility more fully in Sec. VII.

V. OPTICAL CROSS SECTIONS FOR THE ACCEPTOR LEVEL

In this section we report our values for the cross sections σ_p^0 and σ_n^0 corresponding to optical transitions from the valence band to the unoccupied acceptor level and from the occupied acceptor to the conduction band, respectively. Such measurements have been made in the past[3,9,40] using the same techniques as we have used here. In fact, our results on sample *A* are essentially identical to the earlier studies on these types of samples.[3,9] Therefore we will not discuss the technique but only the results. In particular, we will focus mainly on those aspects of the present data which extend the previous work. Specifically, we have made here accurate comparisons of cross-section values between different samples at the same temperature and between different temperatures for the same sample. We have therefore not measured *absolute* cross sections, since this has been done

previously on samples of type A at 90 K.[9]

The optical cross sections for both sample A and B at 80 and 160 K are shown in Figs. 5 and 6. Although the scales of these figures are expressed in arbitrary units, the data for the two samples are on the same scale. Thus the apparent difference in cross-section values of nearly an order of magnitude between samples A and B is a real effect. This strikingly large difference between the samples is quite reproducible and indicates again that the gold-acceptor states in these two types of samples are quite different in spite of their similar thermal-emission rates. In these figures one can also note other less striking differences; namely, the σ_p^0/σ_n^0 ratio for $h\nu$ above 0.8 eV differs in the two samples and the two σ_n^0 curves have slightly different shapes.

The temperature dependence of the cross sections is shown in Fig. 7 for σ_p^0 and in Fig. 8 for σ_n^0. In all cases a shift to longer wavelength and a broadening of the line shape are evident. As can be seen in Fig. 7, the shape and temperature dependence of the hole-emission cross section σ_p^0 is almost exactly the same for the two different samples. The only difference, as mentioned above, is that σ_p^0 in sample B is an order of magnitude larger than in sample A. These data were taken under conditions of identical optical intensity with a defocused monochromator beam providing a uniform illumination. The relative position of the curves was then determined by the time constants of photocurrent and photocapacitance transients. Therefore, the only possible source of error might be a variation in the light intensity *inside* the sample due to differences in such factors as reflectivity, etc. However, since both samples had contacts over only a small fraction of the junction area, we believe that such variations might account for at most only a factor of 2 in the order-of-magnitude discrepancy. Thus we must conclude that, no matter how surprising it might be, the difference between the samples is real.

In contrast to the similar shape and broadening of the σ_p^0 curves for the two different samples, the σ_n^0 data in Fig. 8 indicate slight differences in these properties of the σ_n^0 line shapes. The line shape for sample B shows a slightly more abrupt change of slope near 0.6 eV than does that of sample A. The sample A σ_n^0 line shape is a more uniform curve very similar in shape to the σ_p^0 curves for the two samples. The thermal broadening of the σ_n^0 transition is different in these samples as well. Sample

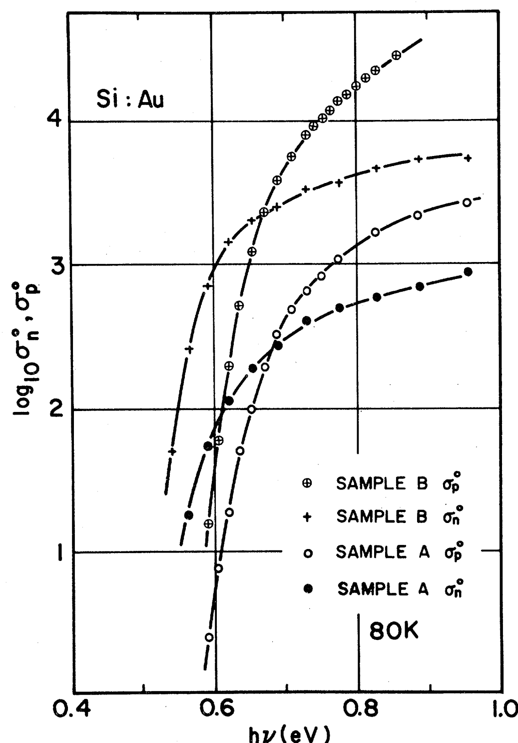

FIG. 5. Optical cross sections for electron and hole emission versus photon energy for gold acceptors at 80 K.

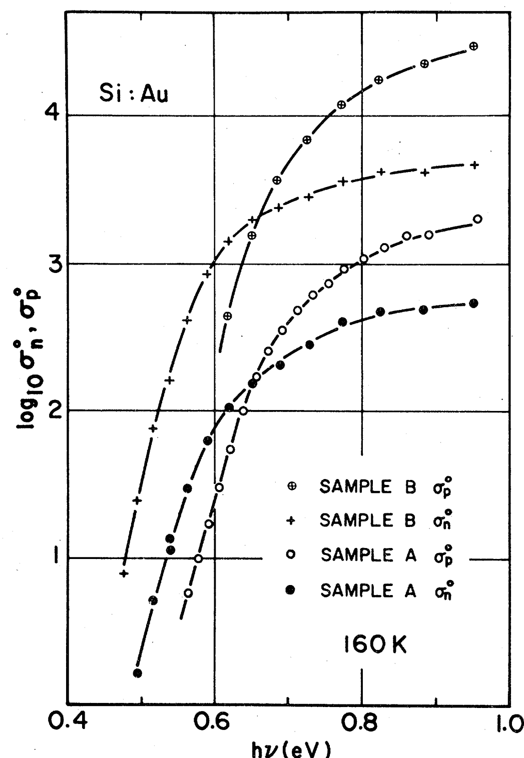

FIG. 6. Optical cross sections for electron and hole emission versus photon energy for gold acceptors at 160 K.

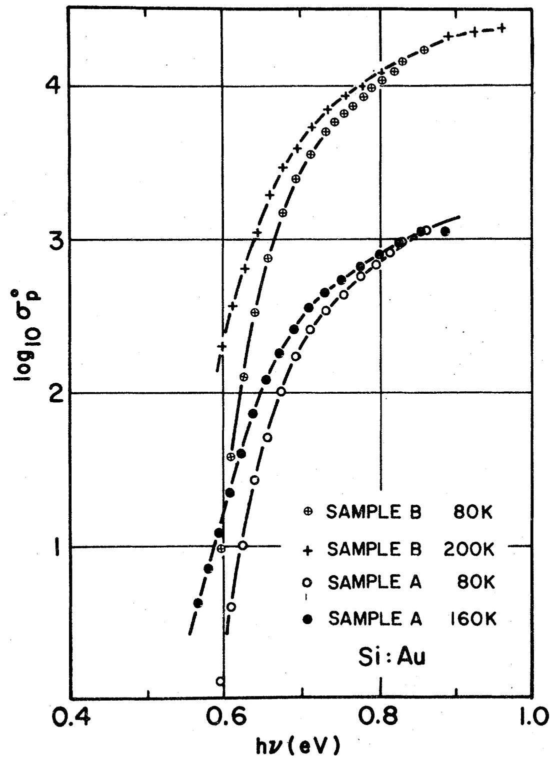

FIG. 7. Optical hole-emission cross sections versus photon energy for gold acceptors showing broadening and shift of threshold for increased temperature.

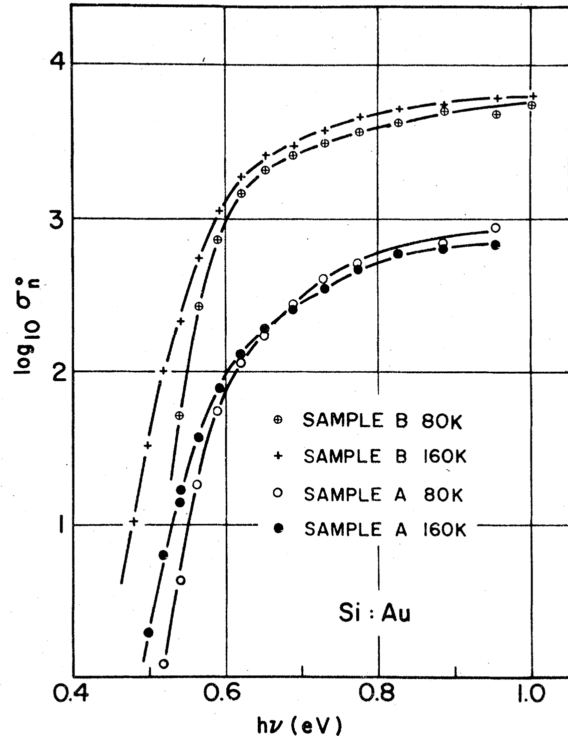

FIG. 8. Optical electron-emission cross sections versus photon energy for gold acceptors showing broadening and shift of threshold with increased temperature.

B has a rather large broadening of σ_n^0, almost exactly the same as the large σ_p^0 broadening in both samples. Sample A, on the other hand, has a somewhat smaller temperature dependence. In fact, over the first decade down from the maximum the shift and/or broadening is barely noticeable, in agreement with the earlier data.[3]

We have not attempted to fit a theoretical line shape to these data but rather have drawn smooth curves through the points. A theoretical fit would clearly be desirable and necessary to extract the threshold energy, temperature shift of threshold, and Franck-Condon (lattice-relaxation) energy. However, since the semiclassical phonon-broadened theory[41] is probably not applicable for the low temperatures and rather modest lattice relaxation apparent from the data, we felt that such a fit would be misleading because of the large number of free parameters which are available in our case. In any event, our main purpose in this paper is to stress the *differences* between two supposedly similar gold centers in two different types of silicon samples. For this it is only necessary to compare the raw data.

Finally, let us compare our present data with the other measurements of the optical properties of the gold acceptor.[3,9,40] First, as we mentioned earlier, our data on sample A agrees with the previously published measurements on the same samples.[3,9] Thus we can use the absolute cross-section values determined earlier for these samples at 90 K.[9] Also in Ref. 9 a comparison is made between samples of type A and the optical data of Tasch and Sah[40] on the gold-acceptor level. Compared to the differences between samples A and B, the differences between sample A and the data of Tasch and Sah are very small. As can be seen in Fig. 6 of Ref. 9, the σ_p^0 values agree almost exactly and the σ_n^0 values differ by, at most, only 50%. Thus our type-A samples and the samples used by Tasch and Sah appear to have a rather similar type of gold defect, whereas our sample B appears to have a quite different defect. It is interesting (and perhaps significant) to note, therefore, that in both sample A and the Tasch and Sah sample the active layer for space-charge spectroscopy was epitaxial n-type silicon whereas in sample B the corresponding n-type layer was Czochralski-grown silicon. We will discuss this point further in Sec. VII.

VI. THERMODYNAMIC ANALYSIS

Recently there has been considerable interest in the temperature dependence of the emission ener-

gies of deep levels.[11] This concept is most conveniently discussed in the language of thermodynamics, as in the recent work of Van Vechten and Thurmond[42] and of Engstrom and Alm.[43] In this section we will briefly discuss the thermodynamic concepts set forth in Refs. 42 and 43 and apply them to the data on the gold-acceptor level in silicon.

A major point of confusion which often arises in discussions of the energy of deep levels is the uncertainty as to which is the proper thermodynamic form of "energy" to use. In a rigorous thermodynamic sense the energy needed to emit one electron from a defect is the chemical potential μ. The chemical potential may in turn be expressed in a number of equivalent ways, two of which are of special interest to us, namely,

$$\mu \equiv \frac{\partial E}{\partial N}\bigg|_{S,V} \equiv \frac{\partial G}{\partial N}\bigg|_{p,T}, \tag{5}$$

where E is the internal energy, G is the Gibbs free energy, N is the number of electrons, S is the entropy, V is the volume, p is the pressure, and T is the absolute temperature. The first definition is relevant to optical transitions while the second applies to thermally stimulated transitions.[42,43] Thus in both cases we can treat the emission energy of a deep level as a Gibbs free energy, defined by

$$G = H - TS, \tag{6}$$

where

$$H = E + pV \tag{7}$$

is the enthalpy. Let us define ΔG_n (ΔG_p) as the free energy needed to emit an electron (hole) from the deep level. This is the proper energy to use in a Boltzmann factor or Fermi function.[43] Then, as shown by Engstrom and Alm,[43] the proper forms for the well-known detailed balance equations relating thermal capture and emission rates are

$$e_n^t = c_n N_c \exp(-\Delta G_n/kT) \tag{8}$$

for electron emission and

$$e_p^t = c_p N_V \exp(-\Delta G_p/kT) \tag{9}$$

for hole emission. In these equations c_n and c_p are the capture coefficients for electrons and holes, respectively, defined by Eq. (4), and N_c (N_V) is the density of states in the conduction (valence) band.

Note that Eqs. (8) and (9) do not include the electronic degeneracy factors in the usual way.[26] This is an important difference. In the proper thermodynamic treatment[43] the changes in degeneracy owing to changes in both the electronic and vibrational states of the system accompanying the emission of an electron or hole are expressed together as the entropy changes ΔS_n or ΔS_p, respectively. Thus from Eqs. (6), (8), and (9) the detailed balance relationships may also be written

$$e_n^t = c_n N_c X_n \exp(-\Delta H_n/kT) \tag{10}$$

for electron emission and

$$e_p^t = c_p N_V X_p \exp(-\Delta H_p/kT) \tag{11}$$

for hole emission. The entropy factors $X_{n(p)}$ are defined as

$$X_{n(p)} = \exp(\Delta S_{n(p)}/k), \tag{12}$$

where $\Delta S_{n(p)}$ is the total entropy change accompanying electron (hole) emission. This total entropy change is in turn equal to the sum of the changes in the entropy due to electronic degeneracy ΔS_{ne} (ΔS_{pe}) and due to atomic vibrational changes ΔS_{na} (ΔS_{pa}). The electronic contribution to the entropy may be expressed in terms of the usual degeneracy factors: g_0 for the defect when unoccupied by an electron, and g_1 for the defect when occupied by a single electron. Thus the entropy factors may be written as

$$X_n = (g_0/g_1)\exp(\Delta S_{na}/k) \tag{13}$$

and

$$X_p = (g_1/g_0)\exp(\Delta S_{pa}/k). \tag{14}$$

From Eqs. (10) and (11) it is clear that the slope of an Arrhenius plot is an enthalpy. If the temperature dependence of $c_n N_c X_n$ (for the example of electron capture and emission) is known, then the "activation energy" of the deep level is ΔH_n, which we shall call the activation enthalpy. If the experimentally determined enthalpy is substituted back into Eq. (10) to calculate c_n without knowing X_n, the resulting value of the capture cross section could be seriously in error, since the entropy factors may be 50 or greater in some cases. This is why in Sec. IV we did not consider capture cross sections obtained in this way. Mircea *et al.*[44] have also pointed out this problem in connection with capture cross-section determinations in GaAs; however, they did not use thermodynamic language.

If the emission rate, capture cross section, and activation enthalpy are all known, then Eqs. (10) or (11) may be used to calculate X_n or X_p. It is not possible, however, to determine from such data the relative magnitudes of the electronic and vibrational parts of the entropy. It is nevertheless of interest to calculate the entropy factors of the gold acceptor in the few cases where sufficient data are known. For our sample A we have $\Delta H_n = 0.553$ eV and at $T = 250$ K; $e_n^t = 10\ \text{sec}^{-1}$, $N_c = 2.39\times10^{19}$ cm^{-3}, and $c_n = 1.29\times10^{-9}\ \text{cm}^3\,\text{sec}^{-1}$. Therefore, $X_n = 48$ or, equivalently, $\Delta S_n = k\ln X_n = 3.87k$. For sample B the thermal-emission rate is essentially identical at several different temperatures but the

capture cross section is 2.4 times larger than in sample A. Therefore $X_n = 20$ or $\Delta S_n = 3.00k$. There are also two cases in the literature where sufficient data are available to calculate X_n, namely, for the sample of Brotherton and Bicknell[4] where we obtain $X_n = 47.5$ ($\Delta S_n = 3.86k$) and for the gold-ion-implanted sample of Nagasawa and Schultz[39] where from their data we calculate $X_n = 1.04$ ($\Delta S_n = 0.04k$). These results are summarized in Table II.

For the case of sample A we are also able to calculate X_p from the previous thermal data on this type of sample.[1] Using the hole-capture cross section determined from the minority-carrier–diffusion length in these samples,[1] we obtain $X_p = 5.7$ at 250 K (sample A), or $\Delta S_p = 1.74k$. It is of interest to compare the total entropy ($\Delta S_p + \Delta S_n$) associated with the creation of an e-h pair via the gold-acceptor level with the entropy of the silicon band gap ΔS_{cv}, i.e., the entropy associated with the direct creation of an e-h pair.[45] There is sufficient data to do this only in sample A where we have $\Delta S_p + \Delta S_n = 5.61k$ at 250 K. From the data compiled by Thurmond,[45] $\Delta S_{cv} = 2.7k$ for silicon at 250 K. Therefore, the entropy to create an e-h pair via the gold-acceptor level apparently is over twice as much as that associated with the direct creation of such a pair. This rather surprising result is not likely to be due to experimental error since the major source of error is in the cross-section data and is on the order of 50% or less. Furthermore, these errors enter only logarithmically into the determination of the entropy.

One of the main consequences of this apparent large total entropy change is that the previous attempts[1,4] to determine the temperature dependence of the acceptor-energy level from thermal data must be reexamined. These earlier analyses implicitly assumed $\Delta S_n + \Delta S_p = \Delta S_{cv}$. Indeed, one would expect this to be true for a simple defect. Engstrom and Grimmeiss[1] analyzed the hole-emission and capture data using $g_0/g_1 = 3.2$. We will restate their results in terms of the thermodynamic language which we have just discussed. From Eq. (14) and the $X_p = 5.7$ value for sample A, we see that $\Delta S_p = 2.9k$, which is very close to $\Delta S_{cv} = 2.7k$ at 250 K. Thus it was argued[1] that the energy level was pinned to the conduction band since the entire entropy of the gap (and hence temperature dependence of the gap) could be accounted for by the hole-emission transition. This follows from the fact that the entropy is a direct measure of the temperature dependence of the Gibbs free energy according to the general thermodynamic relationship.

$$S = \frac{\partial G}{\partial T}\bigg|_{p,T} . \qquad (15)$$

Brotherton and Bicknell,[4] on the other hand, analyzed the thermal capture and emission of electrons at the gold acceptor and came to the opposite conclusion, namely, that the level was pinned to the valence band. In our language, their argument was essentially the following. By using the value of $X_n = 47.5$, which we obtained above from their data, and their value of $g_0/g_1 = 4$ we see from Eq. (13) that $\Delta S_{na} = 2.5k$. Thus if one assumes, as they did, that $\Delta S_n + \Delta S_p = \Delta S_{cv}$, we see that $\Delta S_{pa} = 0.2k$; i.e., the temperature shift of ΔG_p is very small. In fact, however, since $\Delta S_n + \Delta S_p \approx 2\Delta S_{cv}$ we see that the acceptor level seems to move relative to *both* the conduction and valence bands with the total temperature dependence of the gap. Thus both of these earlier studies would seem to be partially correct and partially incorrect.

This analysis leads us, however, to a serious paradox. On the one hand, the data which led above to the large total entropy seem reasonably accurate. On the other hand, according to the data of Ref. 1 and Table II, the free energy of e-h pair production via the gold acceptor in type-A samples at 250 K would seem to be less than the free energy of the gap ΔG_{cv}. This violates the conservation of energy. The value of $\Delta G_n = \Delta H_n - T\Delta S_n$ is given in Table II. From the data of Ref. 1 we showed above that $\Delta S_p = 1.74k$. Thus taking ΔH_p (250 K) $= 0.641$ eV,[1] we see that ΔG_p (250 K) $= 0.604$ eV. The total enthalpy ($\Delta H_n + \Delta H_p$) is equal to 1.194 eV, which is remarkably close to ΔH_{cv} (250 K) $= 1.194$.[45]

TABLE II. Summary of the thermodynamic parameters for electron capture and emission at the gold acceptor.

Sample	e_n^t (250 K) (sec^{-1})	σ_n (250 K) (10^{-16} cm^2)	ΔH_n (eV)	X_n	$\Delta S_n/k$	ΔG_n (250 K) (eV)
A	10	0.69	0.553	48	3.87	0.470
B	10	1.7	0.553	20	3.00	0.488
Brotherton and Bicknell (Ref. 4)	10	0.9	0.555	47.5	3.86	0.472
Nagasawa and Schulz (Ref. 38)	8.5	21.8	0.54	1.04	0.04	0.54

The total Gibbs free energy $(\Delta G_n + \Delta G_p)$, however, is only 1.074 eV at 250 K. This is 63 meV *less* than the free energy of the gap [ΔG_{cv} (250 K) = 1.137 eV] at 250 K.[45] This cannot possibly be correct. Therefore we must conclude that the total enthalpy $(\Delta H_n + \Delta H_p)$ is too low or that the total entropy $(\Delta S_n + \Delta S_p)$ is too high. The former is not very likely, however, since all measurements of $\Delta H_n + \Delta H_p$ have been less than or equal to ΔH_{cv} at the temperature of measurement.[1,16,18,24,40] Therefore we are left with the possibility that the value of $\Delta S_n + \Delta S_p \approx 2\Delta S_{cv}$ may be incorrect. The origin of any possible errors, however, is not at all obvious.

One possible problem is that all of the above analysis is based on the implicit assumption that it is valid to use the detailed balance relationships to determine $X_{n(p)}$ from the capture and emission data. This is certainly true if both the capture and emission rates are measured in neutral material. In the case of space-charge spectroscopy, however, the emission rate is measured in the presence of the rather large electric field which is always present in the space-charge layer. The capture cross sections, on the other hand, are usually measured in zero field. Thus even though we have shown that the emission rate e_n^t is practically independent of applied bias, it still may be true that rather small electric fields (smaller than the smallest field at the gold defects in zero-applied bias) have a significant enough effect on the emission rate to invalidate the use of Eqs. (8)–(11) in the present case.

It also may be possible that the capture occurs via excited states of the gold level in such a way as to modify the form of the detailed balance relationship. Capture via an excited state gives rise to a characteristic temperature dependence of the electron-capture rate corresponding to thermal reemission from the excited state occurring before deexcitation to the ground state.[7,46,47] Thus the capture cross section *increases* with decreasing temperature with a negative activation energy equal to the depth of the excited state. Although most measurements indicate a temperature-independent σ_n, there is some very slight evidence for a small increase in σ_n at lower temperatures with an activation energy of −2.6 meV.[36] There is no evidence in the capture data for any deeper excited states. Therefore, if such a state were present it would have to be of such a nature that the rate of deexcitation following capture was sufficiently rapid that thermal reemission could not occur. Such a state would not affect the temperature dependence of the capture rate and would not alter the measured activation enthalpy in zero field. However, if an electric field were present which enhanced the emission process from the excited state so that its rate now exceeded the rate of deexcitation, then the apparent activation enthalpy would be reduced by the depth of the excited state.[47] If this were the case, then the total enthalpy would *exceed* the measured sum $\Delta H_n + \Delta H_p$ by the depth of the excited state. Therefore, the true total enthalpy for the gold-acceptor case would exceed ΔH_{cv} and a large total entropy would not violate the conservation of energy by making possible the creation of e-h pairs via the gold level with $\Delta G_n + \Delta G_p < \Delta G_{cv}$. A problem with this idea, however, is that the ΔH_n is very similar in both junction and Hall-effect measurements.[1]

Another possible explanation for the large entropy of e-h pair production via the gold acceptor has been suggested by Van Vechten.[48] This is based on his original proposal[42] that "Au_{Si}" is actually a complex of an Au interstitial with a vacancy, $Au_i V_{Si}$. Thus the ionization reactions for the gold acceptor Au_{Si} could actually be

$$Au_i V_{Si} \rightarrow Au_i V_{Si}^- + e^+ \quad (16)$$

for the hole emission, and

$$Au_i V_{Si}^- \rightarrow Au_i^+ V_{Si}^- + e^- \quad (17)$$

for electron emission. Each of these reactions would be expected to have the entropy of the gap ΔS_{cv}, so that the total entropy of e-h pair production via the $Au_i V_{Si}$ defect would be $2\Delta S_{cv}$ as is observed. The apparent nonconservation of energy could also be explained if the final $Au_i^+ V_{Si}^-$ state had lower energy than the initial $Au_i V_{Si}$ state. Thus the Au defect could not be returned to its initial state without some additional energy being supplied to drive the internal defect reaction $Au_i^+ V_{Si}^- \rightarrow Au_i V_{Si}$. This internal reaction energy must be equal to or greater than the energy deficit $\Delta G_{cv} - (\Delta G_n + \Delta G_p) = 63$ meV (at 250 K).

Without further data, however, these and other possible explanations for the paradox of the apparently large total entropy have to be regarded as speculative. One point is clear, however. As is also evidenced by the other data in this paper, the gold acceptor is not the simple defect it was once thought to be.

VII. DISCUSSION OF GOLD-DEFECT MODELS

The results which we have presented on gold-doped silicon leave one with a rather confused picture of the true nature of the gold-related deep levels in this material. In this section we will discuss some of the main features of the data and attempt to bring some order to the apparent chaos by proposing several speculative defect models which might possibly explain certain aspects of the data.

One of our major assumptions throughout the paper has been that the data obtained by different people using different techniques on different samples are basically correct. Therefore one cannot appeal to unspecified experimental errors to explain some of the rather large discrepancies which have been reported. This assumption is based on the data presented in this paper which verify that there are real differences between the gold-related acceptor levels in different types of silicon samples.

There are four major points from our data and those in the literature which any microscopic model must be able to explain:

(1) The gold donor and acceptor levels seem to be related to *different* defects which are often present in roughly equal concentrations.

(2) The electron thermal-emission properties of the Au-acceptor defect are almost identical to the acceptors in Ag-, Co-, and Rh-doped silicon and very similar to the deep donors in S-doped and quenched silicon.

(3) Differences of a factor over 30 in the electron-capture cross section of the gold acceptor can be correlated with the concentration of gold relative to that of the shallow donor.

(4) The optical cross sections of gold-acceptor defects in an n-type Czochralski-grown silicon sample are approximately ten times larger than the corresponding cross sections in an n-type epitaxial silicon sample.

Our data relate primarily to the acceptor defect. Therefore, we will discuss only models for this deep level. The striking similarity of the e_n^t vs T behavior for six different defects (Au, Ag, Co, Rh, S, and quenched-in) suggests that the underlying structure of these defects may be very similar in all cases. Thus the electron thermal-emission rate must be very insensitive to the chemical nature of the associated impurity and even to whether or not the net behavior of the complex is as a donor or an acceptor. The fact that the quenched-in midgap level is among these six similar defects may give us an important clue as to the common structure. Sah and Wang[20] have shown that the quenched-in defects are generated in a strained or disordered surface layer during heat treatment and diffuse rapidly into the bulk of the sample above 1100°C. Samples with chemically polished surfaces do not generate such defects.[20] It was argued, therefore, that these quenched-in defects are most likely vacancy complexes.[20] They must not be simple vacancies or divancies, however, since the 0.55-eV electron-emission signal is not seen in (1–10)-MeV electron-irradiated silicon.[49]

It is tempting to expand upon the ideas of Van Vechten[42] and to speculate that the quenched-in vacancy complex is the basic element of the $E_c - 0.55$ eV family of defects. The various chemical impurities might then be able to associate with this vacancy complex without significantly affecting its thermal-emission properties. Since such centers are typically not present in control samples which have had the same heat treatment as the impurity-diffused samples, it would be necessary for the diffusion process itself to generate or stabilize these vacancy complexes. Indeed, there is evidence that vacancy-related defects diffusing in from the surface control the transition from the mobile gold interstitial to the final gold defect state during gold diffusion in silicon.[50–52] It makes intuitive sense that a large atom such as gold would most readily be incorporated into the silicon lattice if a large "hole" in the form of a vacancy complex were available to it. There is also recent evidence[53] that gold atoms have an enhanced solubility (i.e., are "gettered") near dislocations in silicon. This implies that the gold atom would prefer to occupy a distorted location in the lattice. From a theoretical point of view one would also expect vacancy-related defects to behave very much like the underlying vacancy part of the complex even in the presence of associated chemical impurities.[54] Indeed, calculations by Hemstreet[55] of "substitutional" transition elements in Si suggest that elements to the left of Co in the periodic table are bonded to their Si neighbors while Co, Ni, Cu, and Zn maintain their d-like atomic character with the associated deep states in the gap being due to the broken bonds on the Si neighbors, i.e., they are vacancylike. Similar results have recently been obtained in much more detail for Zn in Si by Bernholc, Lipari, and Pantelides using totally different theoretical techniques.[56]

The tendency of gold and other heavy metal impurities in silicon to have an enhanced solubility in regions of large phosphorous (donor) concentration is well known.[57,58] For example, in regions where the P concentration is greater than 10^{20} cm^{-3}, the substitutional Au concentration is 2×10^{18} cm^{-3} after a 100-h heat treatment at 800°C.[53] This is to be contrasted with the normal gold solubility limit of less than 10^{15} cm^{-3} at 800°C.[24] This enhanced solubility or gettering effect is evidence for a tendency to form ion pairs between phosphorous donors and gold acceptors. Gold-diffusion data in heavily doped silicon also support the idea that gold and phosphorus tend to form complexes.[59] Such a pairing effect may also explain the large differences in electron-capture cross section in Table I between samples where the gold concentration is either much less than or much greater than the shallow donor (presumably phosphorous) concentration. In the $N_{Au} \gg N_D$ case there is little possibility for donor-acceptor pairs to form.

Thus, these types of samples should correspond to a gold-acceptor complex which is not paired with a shallow donor impurity. In the samples with $N_{Au} \ll N_D$, on the other hand, it is possible that a significant number of the gold acceptors are associated with donor atoms. This type of complex between a metallic impurity and a donor has recently been seen in Cu-doped GaP.[8] Also, in Co-doped Si several different Co-related defect complexes have been seen by Mössbauer measurements in samples with P or B concentrations greater than about 10^{16} cm^{-3}.[60]

Among the samples in the $N_{Au} \ll N_D$ regime we find our final set of discrepancies, namely, the optical and capture cross-section differences between samples A and B. As we have noted previously, sample A is epitaxial silicon and sample B is Czochralski silicon. One of the major differences between these two kinds of silicon is that Czochralski silicon has a rather large concentration of oxygen ($>5\times10^{17}$ cm^{-3}) whereas epitaxial silicon typically has very little oxygen ($<10^{15}$ cm^{-3}). Since the epitaxial sample A has optical properties close to another epitaxial sample in the literature,[40] we suspect that it is the Czochrolski sample B which is unusual, perhaps because of oxygen impurities associated with the gold complex.

We believe, therefore, that the various discrepancies in our data and in the literature can be taken as evidence that the deep levels in gold-doped silicon are far more complicated than had previously been assumed. Although the microscopic structures of the various gold complexes are by no means obvious, we think the speculative models which we have discussed may be helpful in focusing attention on several possible areas of future inquiry. These particular models may ultimately be proven right or wrong. The major point is to recognize that the deep levels due to gold in silicon must be associated with some sort of complex structure.

VIII. SUMMARY AND CONCLUSIONS

The major conclusion of this study is that the deep levels in gold-doped silicon are not due to a simple substitutional gold atom with three charge states, as had almost universally been assumed. We find that the donor and acceptor levels in our samples are not related to the same gold center but are due to different gold-related defects which occur in roughly equal concentrations in most samples, but are unequal in others.

Our major emphasis has been to study carefully the gold-acceptor defect in two different types of silicon p^+n diodes: one type fabricated from epitaxial silicon (sample A) and another from Czochrolski silicon (sample B). The electron thermal-emission rates (and hence DLTS peak positions) of the gold acceptor in these two types of samples are the same within experimental error. Indeed, as we point out, it is quite surprising to note that at least five other deep levels in silicon (Ag, Co, Rh, S, and process-induced defects) have electron thermal-emission rates and activation energies which are indistinguishable from the gold acceptor. We suggest, therefore, that the quenched-in centers induced by heat treatment may be the basic defect of this family.

In contrast to the electron thermal-emission properties, however, we find that the thermal capture and optical cross sections of the gold acceptor show quite a wide range of variation both between our two types of samples and between our samples and data in the literature. We show that the differences between samples A and B are not due to systematic errors in the various measurement techniques. If we further assume that the data in the literature are also basically reliable, we can show that the widely divergent values reported for the electron-capture cross section of the gold acceptor can be correlated with the ratio of the gold concentration to that of the shallow donor. We propose, therefore, that the variations in thermal-capture cross section are related to the formation of complexes between gold atoms and shallow donors (such as phosphorous). We also find that the gold-donor defect in our n-type samples has very different thermal-capture cross sections from those reported in the literature for gold donors in p-type silicon. Thus, the gold donor may also be forming complexes with the shallow dopants.

The origin of the large differences in the magnitudes of the optical cross sections of the acceptor in our two types of samples and the somewhat smaller differences between the shapes of these cross sections as a function of photon energy are not completely understood. We note, however, that the data on our epitaxial samples are consistent with data in the literature for other epitaxial samples. Therefore, we propose that the origin of the differences may be related to the different methods of crystal growth between our two types of samples (epitaxial versus Czochrolski) and perhaps may be related to the large differences in oxygen concentration which are inherent to these growth methods.

Finally, we have attempted to extract the temperature dependence of the gold-acceptor level from our data. We show that the thermal data are consistent with the literature and can lead one to the conclusion that the acceptor level is fixed either to the conduction band or to the valence band, depending on the method of analysis and on one's initial assumptions. The optical data are difficult to in-

terpret due to the fact that the thermal broadening obscures any shifts in the threshold energy with temperature.

Thus we conclude that the "gold acceptor" is actually a family of closely related defects which are much more complicated than simple substitutional Shockley-Read-Hall deep levels. Further study is clearly needed to elucidate the underlying energy-level structures as well as the general spatial and chemical structures of this family of gold defects.

ACKNOWLEDGMENTS

Two of the authors (H.G.G. and E.M.) would like to thank the Swedish Natural Science Research Council and the Swedish Board for Technical Development for their financial support. One of us (D.V.L.) would also like to thank these sources as well as the staff and students of the Department of Solid State Physics of the Lund Institute of Technology for their support and hospitality during a visit to Lund. We also gratefully acknowledge many stimulating discussions with S. Braun, O. Engstrom, E. Janzen, L.-A. Ledebo, C. Ovren, and B. Skarstam.

APPENDIX: CARRIER CAPTURE AND SATURATION OF THE GOLD DONOR

In Sec. III we interpreted the DLTS spectra in Fig. 3 as showing that the gold donor had a lower concentration than the gold acceptor in sample B, and hence that the two levels must relate to different defects. In order for this interpretation to be valid we must be able to show that the DLTS signals for the donor and the acceptor are in fact proportional to their respective concentrations. For a majority-carrier trap such as the gold acceptor in n-type Si this is fairly straightforward. One must simply show that the length of the majority-carrier pulse is sufficient to totally fill the trap.

For a minority-carrier trap such as the gold donor in n-type Si, on the other hand, the problems of measuring concentration are more severe. The first requirement that must be met is that the injected current pulse used to fill the trap with minority carriers is sufficiently intense so that the minority-carrier–capture rate is much greater than that for majority carriers. The second requirement is that the pulse be long enough to reach steady-state conditions. If these two conditions are met one says that the trap is "saturated" with minority carriers. Even then, however, it is possible that the resulting DLTS signal may not be a true measure of the defect concentration. This may occur if the fall time of the injection pulse is on the order of or longer than the majority-carrier–capture time. Thus, for example, in the case of the gold donor in n-type silicon one might be able to saturate the level with holes during a sufficiently long and intense forward-bias pulse but the level might refill with electrons during the time it takes the pulse to decay from zero volts to the steady reverse-bias voltage. In our samples such an effect will occur for switching times on the order of a few tenths of a microsecond. Thus rather fast fall times are necessary to obtain a true DLTS concentration signal for the gold donor in p^+n junctions. As we will show, the donor signals in our samples (Fig. 3) satisfy all of these requirements.

The proper analysis of minority-carrier capture and trap saturation has been discussed by Henry *et al.*[14] They showed from a straightforward analysis of the relevant rate equation for carrier capture at a minority-carrier trap that the time constant of the trapped minority-carrier concentration during a forward-bias injection pulse was

$$\tau^{-1} = c_n n + c_p p \,, \tag{A1}$$

where n is the free electron concentration, p is the hole concentration, and c_n and c_p are defined as in Eq. (4). In this analysis we will assume that the temperature is low enough so that thermal-emission rates may be neglected.

For the case of the gold donor in n-type silicon n is fixed by the net donor concentration and p is related to the injected current I by

$$p = (\gamma L_p / qAD_p) I \,, \tag{A2}$$

where γ is the injection efficiency, D_p is the hole-diffusion coefficient, L_p is the hole-diffusion length, q is the electronic charge, and A is the diode area. The steady-state trapped-hole occupation p_T on the donor level is

$$\frac{p_T}{N_{\mathrm{Au}(D)}} = \frac{c_p p}{c_n n + c_p p} \,, \tag{A3}$$

where $N_{\mathrm{Au}(D)}$ is the concentration of gold-donor defects. Note that $p_T = N_{\mathrm{Au}(D)}$ only when the injected current (and hence p) is large enough so that $c_p p \gg c_n n$. This is the so-called saturation condition. If the fall time of the injection pulse is fast enough, the trapped-hole occupation is then proportional to the DLTS signal.

The saturation condition can be experimentally recognized in two ways. The first is to increase I until the DLTS signal reaches its maximum value with no further increase for larger I [see Eq. (A3)]. The second is to compare the rate of the buildup of the DLTS signal $(c_n n + c_p p)$ with the electron-capture rate $(c_n n)$ measured using the second pulse of a two-pulse injection-clear sequence.[61] This two-pulse scheme for measuring the majority-carrier–capture cross section of a minority-car-

rier trap is also an excellent way to verify that the trapped holes are not being lost by electron capture during the fall time of the injection pulse. This can easily be done by directly comparing the measured electron-capture time constant with the fall time of the injection pulse. In fact, the two-pulse sequence is essentially the same as a direct measurement of the DLTS-signal decrease as a function of the fall-time of the injection pulse. By making measurements of these sorts we can thus show that the DLTS hole-emission spectra of the gold donor in Fig. 3 correspond to a saturating injection pulse with negligible loss of signal during the 10-nsec fall time of the pulse. Thus the DLTS spectra are a true measure of the concentrations of the donor and acceptor defects.

In view of the widely accepted belief that gold is a three-charge-state system, one might ask whether it could be possible that the apparent inequality of the donor and acceptor DLTS signals is due to a peculiarity of the coupled rate equations for a three-state system and not the result of two different defects of different concentrations. We have examined this point in detail using the coupled equations of Ref. 14. We find that in our case the two levels should behave independently even if they were coupled. This implies that the donor and acceptor signals must be equal (except for possible systematic differences due to the edge region of the space-charge layer) if the Si:Au system is to be considered a coupled three-state defect such as O in GaP.[14]

The key observation which supports the assertion that the levels behave independently even in the context of the three-state model is that the acceptor level saturates with holes more readily than does the donor, i.e., at an injected current level for which the DLTS electron-emission signal from the acceptor disappears (as in Fig. 3) the DLTS hole-emission signal from the donor is not yet fully developed. If gold were a coupled three-state system and were to show a spurious and systematic reduction in the apparent concentration of the one-electron state (neutral Au) relative to the two-electron state (ionized Au acceptor in the three-state model), the gold-acceptor DLTS signal would not be reduced to zero by an injection pulse which apparently saturated the donor. In this case even if a maximum number of gold centers were in the one-electron (neutral) state, some fraction of these defects would be occupied by two electrons and hence would not give rise to a donor DLTS signal.

Perhaps a simpler and more convincing argument that coupled rate equations could not explain our results, however, is the fact that the donor-to-acceptor ratio is different in our different samples and has been reported by Engstrom[27] to change in the *same* sample during low-temperature heat treatment. A systematic effect of the coupled rate equations would give rise to a fixed donor-to-acceptor ratio which should be an inherent property of the defect and should not vary in samples of different types or with different thermal histories.

In the process of verifying the saturation conditions, we have also been able to measure as a side product for the first time the capture cross sections of the gold-donor defect in n-type material. The results are quite different from those reported in the literature from measurements made on p-type samples. The measurements of capture cross sections for the gold-donor defect are not as extensive as for the acceptor (see Table I), but there is general agreement among the various results.[22,25,35] The cross section for the capture of an electron on the positive gold donor in p-type silicon has been measured by both minority-carrier lifetime[25] and photoconductive-decay[35] methods with an average value of $\sigma_n = 3.3\times10^{-15}$ cm^2 at 300 K and a $T^{-2.5}$ temperature dependence between 200 and 500 K. The capture of a hole on the neutral gold donor in p-type silicon was found to be $\sigma_p = 1.6\times10^{-15}$ cm^2 at 135 K from DLTS measurements[22] and $\sigma_p = 1.4\times10^{-15}$ cm^2 at 300 K from photoconductive-decay measurements.[35] Thus there is apparently little temperature dependence of this cross section, as would be expected for a neutral center with such a large cross section.[7]

Note that these results in the literature would imply that the donor defect could not be saturated with holes in n-type material because $\sigma_n > \sigma_p$. In fact, at the temperature of the donor DLTS peak in Fig. 3 we would expect from the $T^{-2.5}$ extrapolation that $\sigma_n \approx 1.6\times10^{-14}$ cm^2. Thus $\sigma_n \approx 10\sigma_p$. With these cross sections the donor-hole-emission signal in Eq. (A3) would be only 10% of the true donor concentration even under very high injection conditions where $n = p$. Furthermore, the time constant for electron capture would be on the order of 1 nsec in our samples and thus much less than our injection-pulse fall time of 10 nsec. Therefore, the gold-donor defects reported in p-type material could not have been seen by DLTS in our p^+n samples even if they were present. Such centers could have been seen by DLTS in p-type material, however, where they would be majority-carrier traps. They could also have been seen by photocapacitance or photocurrent measurements in our samples. Therefore since the photocurrent measurements of Braun and Grimmeiss[9] also showed equal donor and acceptor concentrations in the epitaxial samples of type A, we must conclude that only donor levels of the type seen by DLTS in Fig. 3 are present in these samples. This is quite significant since DLTS alone would be unable to rule out the presence of the p-

type donor.

The very fact that we can see DLTS signals due to donor levels in our n-type samples immediately implies that the capture cross sections must be different from those reported in the literature. At 158 K we find $\sigma_n = 4\times10^{-16}$ cm^2 in sample B and $\sigma_n = 1-2\times10^{-16}$ cm^2 in sample A. This is about a factor of 50 less than the extrapolated electron-capture cross section at 158 K which has been measured in p-type silicon. The hole-capture cross section in our samples is more difficult to measure but can be estimated to be on the order of 10^{-14} cm^2. This is nearly an order of magnitude larger than in p-type silicon. Thus in our samples, we have $\sigma_p \gg \sigma_n$ so that saturation is readily attainable.

One might wonder why the capture cross sections of the gold-donor defects are so different in p- and n-type silicon when the thermal-emission properties are so similar. This is similar to the situation we have discussed in Sec. IV for the gold-acceptor defects. Perhaps the donor defect is also a complex between gold atoms and donor or acceptor atoms where the nature of this complex affects the capture cross sections more strongly than it does the thermal-emission rate.

We should note here that others have also seen several different donors in Au-doped Si. In addition to the observations of Engstrom mentioned earlier, Scheibe and Schroter[62] have reported seeing three different donors in the lower half of the gap in Au-doped Si. One of these was thought to be due to the heat treatment and could be annealed out at 260°C. However, two donors which were clearly gold-related were seen as well. One at $E_V + 0.4$ eV annealed out below 400°C while the other at $E_V + 0.35$ eV annealed at 580°C. If the capture cross section of the $E_V + 0.4$ eV donor were an order of magnitude greater than that for the $E_V + 0.35$ eV level, then the DLTS-peak positions under the conditions of Fig. 3 would be difficult to distinguish. It is therefore possible that the donor level seen by Engstrom and by us is the $E_V + 0.4$ eV level seen by Scheibe and Schroter.

[1]O. Engstrom and H. G. Grimmeiss, J. Appl. Phys. 46, 831 (1975).
[2]R. Kassing and H. Lenz, Phys. Status Solidi A 25, 131 (1974).
[3]O. Engstrom and H. G. Grimmeiss, Appl. Phys. Lett. 25, 413 (1974).
[4]S. D. Brotherton and J. Bicknell, J. Appl. Phys. 49, 667 (1978).
[5]D. V. Lang and M. Jaros (unpublished).
[6]E. P. Sinyavskii and V. A. Kovarskii, Fiz. Tverd. Tela. (Leningrad) 9, 1464 (1966), [Sov. Phys—Solid State 9, 1142 (1967)].
[7]C. H. Henry and D. V. Lang, Phys. Rev. B 15, 989 (1977).
[8]P. O. Fagerstrom, H. G. Grimmeiss, and H. Titze, J. Appl. Phys. 49, 3341 (1978).
[9]S. Braun and H. G. Grimmeiss, J. Appl. Phys. 45, 2658 (1974).
[10]S. Braun and H. G. Grimmeiss, J. Appl. Phys. 44, 2789 (1973).
[11]H. G. Grimmeiss, Ann. Rev. Mater. Sci. 7, 341 (1977).
[12]G. L. Miller, D. V. Lang, and L. C. Kimerling, Ann. Rev. Mater. Sci. 7, 377 (1977).
[13]H. Kukimoto, C. H. Henry, and F. R. Merritt, Phys. Rev. B 7, 2486 (1973).
[14]C. H. Henry, H. Kukimoto, G. L. Miller, and F. R. Merritt, Phys. Rev. B 7, 2499 (1973).
[15]A. Zylbersztejn, Appl. Phys. Lett. 33, 200 (1978).
[16]C. T. Sah, L. Forbes, L. I. Rosier, A. F. Tasch, Jr., and A. B. Tole, Appl. Phys. Lett. 15, 145 (1969).
[17]L. D. Yau and C. T. Sah, Appl. Phys. Lett. 21, 157 (1972).
[18]L. C. Parrillo and W. C. Johnson, Appl. Phys. Lett. 20, 104 (1972).
[19]L. D. Yau and C. T. Sah, Solid-State Electron. 17, 193 (1974).
[20]C. T. Sah and C. T. Wang, J. Appl. Phys. 46, 1767 (1975).
[21]L. L. Rosier and C. T. Sah, Solid-State Electron. 14, 41 (1971).
[22]J. A. Pals, Solid-State Electron. 17, 1139 (1974).
[23]D. V. Lang, J. Appl. Phys. 45, 3023 (1974).
[24]C. B. Collins, R. O. Carlson, and C. J. Gallagher, Phys. Rev. 105, 1168 (1957).
[25]G. Bemski, Phys. Rev. 111, 1515 (1958).
[26]A. G. Milnes, *Deep Impurities in Semiconductors* (Wiley, New York, 1973).
[27]O. Engstrom (private communication).
[28]F. Willmann, M. Blatte, H. J. Queisser, and J. Treusch, Solid State Commun. 9, 2281 (1971).
[29]F. Willmann, D. Bimberg, and M. Blatte, Phys. Rev. B 7, 2473 (1973).
[30]D. V. Lang and R. A. Logan, J. Electron. Mater. 4, 1053 (1975).
[31]D. V. Lang and R. A. Logan (private communication).
[32]H. D. Barber, Solid-State Electron. 10, 1039 (1967).
[33]R. R. Senechal and J. Basinski, J. Appl. Phys. 39, 3723 (1968).
[34]Y. Zohta, J. Appl. Phys. 43, 1713 (1972).
[35]J. M. Fairfield and B. V. Gokhale, Solid-State Electron. 8, 685 (1965).
[36]J. Barbolla, M. Puguet, J. C. Brabant, and M. Brousseau, Phys. Status Solidi A 36, 495 (1976).
[37]R. Kassing, E. Kohler, and I. Dudeck, Phys. Status Solidi A 30, 141 (1975).
[38]W. D. Davis, Phys. Rev. 114, 1006 (1959).
[39]K. Nagasawa and M. Schulz, Appl. Phys. 8, 35 (1975).
[40]A. F. Tasch, Jr. and C. T. Sah, Phys. Rev. B 1, 800 (1970).
[41]M. Jaros, Phys. Rev. B 16, 3694 (1978).
[42]J. A. Van Vechten and C. D. Thurmond, Phys. Rev. B 14, 3539 (1976).

[43]O. Engstrom and A. Alm, Solid-State Electron. 21, 1571 (1978).
[44]A. Mircea, A. Mitonneau, and J. Vannimenus, J. Phys. (Paris) Lett. 38, L-41 (1977).
[45]C. D. Thurmond, J. Electrochem. Soc. 122, 1133 (1975).
[46]R. M. Gibb, G. J. Rees, B. W. Thomas, B. L. H. Wilson, B. Hamilton, D. R. Wight, and N. F. Mott, Philos. Mag. 36, 1021 (1977).
[47]D. V. Lang (unpublished).
[48]J. A. Van Vechten (private communication).
[49]L. C. Kimerling, in *Radiation Effects in Semiconductors, 1976,* edited by N. B. Urli and J. W. Corbett (Institute of Physics, London, 1977), p. 221.
[50]W. M. Bullis, Solid-State Electron. 9, 143 (1966).
[51]M. Yoshida and K. Saito, Jpn. J. Appl. Phys. 9, 1217 (1970).
[52]F. A. Huntley and A. F. W. Willoughby, Solid-State Electron. 13, 1231 (1970).
[53]W. F. Tseng, T. Koji, J. W. Mayer, and T. E. Seidel, Appl. Phys. Lett. 33, 442 (1978).
[54]M. Jaros and S. Brand, Phys. Rev. B 14, 4494 (1976).
[55]L. A. Hemstreet, Phys. Rev. B 15, 834 (1977).
[56]J. Bernholc, N. O. Lipari, and S. T. Pantelides, Phys. Rev. B 21, 3545 (1980).
[57]R. L. Meek and C. F. Gibbon, J. Electrochem. Soc. 121, 444 (1974).
[58]T. E. Seidel and R. L. Meek, in *Ion Implantation in Semiconductors and Other Materials*, edited by B. L. Crowder (Plenum, New York, 1973), p. 305.
[59]R. Sh. Malkovich and V. A. Poloeva, Fiz. Tverd. Tela (Leningrad) 18, 2606 (1976) [Sov. Phys.—Solid State 18, 1521 (1976)].
[60]W. Schroter (private communication).
[61]D. V. Lang, J. Appl. Phys. 45, 3014 (1974).
[62]E. Scheibe and W. Schroter, in *Radiation Effects in Semiconductors, 1976,* edited by N. B. Urli and J. W. Corbett (Institute of Physics, London, 1977), p. 221.