Appl. Phys. A 48, 3–9 (1989)

Applied Physics A Solids and Surfaces

Introduction to Defect Bistability

A. Chantre

CNET-CNS, BP 98, F-38243 Meylan Cedex, France

Received 16 June 1988/Accepted 2 September 1988

Abstract. Molecule-like defects in crystalline semiconductors can display bistable properties when two different structural arrangements are possible for constant charge on the defects. This introductory paper will review the physical and technical aspects of this new field of research. It also includes a detailed presentation of our experimental studies on the recently discovered bistable thermal donor defect in silicon.

PACS: 61.70, 71

Reports of defects in semiconductors exhibiting bistable electronic properties have become more and more frequent in the recent literature. After the pioneering compound semiconductor studies, silicon studies have emerged and progressed even faster [1]. Most important in this regard have been the microscopic identifications of two bistable defects in crystalline silicon: the iron-aluminum [2] and carbon-carbon [3] pairs. Today, physical mechanisms of defect bistability have been elucidated. At the same time, techniques to isolate and fully characterize these defects are well established.

This paper is intended to provide an introduction for those people becoming interested in this new field of research. It is divided in two main sections. In Sect. 1, we first introduce the concept of defect bistability with the help of configuration coordinate diagrams. Then we review the experimental techniques and procedures which have been developed to study these defect systems. This is followed by a discussion of the possible physical origins of bistability. Section 2 contains a detailed presentation of our latest investigations on a technologically important bistable center in silicon, the so-called thermal donor defect.

1. Bistable Defects: Physics and Characterization

1.1. Bistability, Multistability and Metastability

It is convenient to use configuration coordinate (C.C.) diagrams to introduce defect bistability. In these diagrams, the total (electronic plus elastic) energy of a defect is plotted versus some configuration coordinate for its different charge states.

For usual, monatomic point-like semiconductor defects, only a single structural configuration is possible for any given charge state. In the simplest situations (e.g. substitutional dopant impurities in Si), this configuration is the same for all charge states. A C.C. representation of such an ordinary defect is shown in Fig. 1a. There are also more complex situations where a change in defect charge state is accompanied by significant structural distortions around it. The familiar picture of displaced C.C. parabola is then used (Fig. 1b). A prototypical example of such defects is the famous DX center in AlGaAs. Let us stress again that, even for these so-called large lattice relaxation defects,

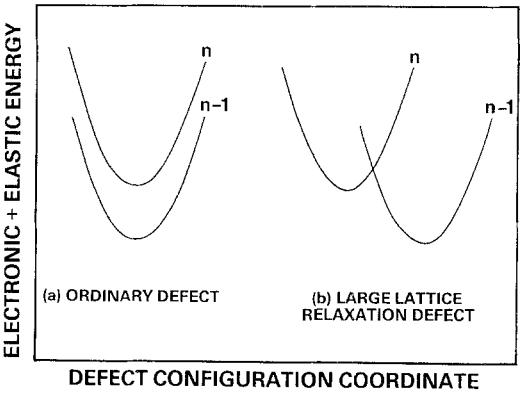

Fig. 1. Configuration coordinate diagrams for usual semiconductor defects

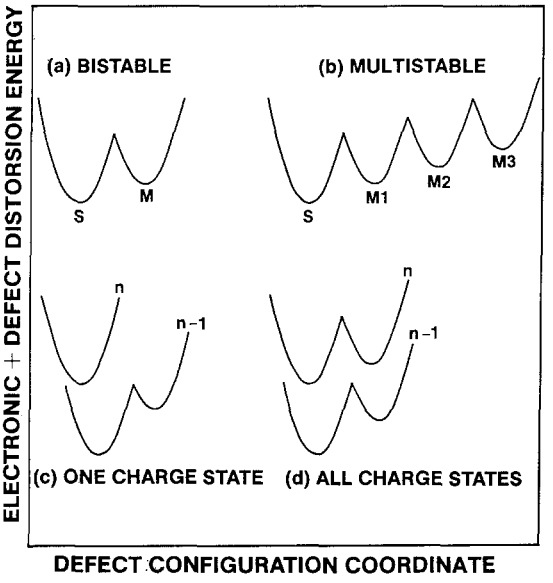

Fig. 2. Potential surfaces for different classes of "metastable" defects in semiconductors

there is a *unique* atomic arrangement available for any charge state of the defects.

In contrast, *bistability* refers to the situation where a defect can exist *in two different structural configurations for the same charge state.* Of course, one of these configurations is metastable (M), and separated from the stable one (S) by a barrier (Fig. 2a). Multistability – the existence of more than one metastable configuration (Fig. 2b) – has also been demonstrated recently for a radiation damage defect in Si [4], but will not be discussed further in this paper. The EL2 defect in GaAs, and the bistable TD defects in Si (Sect. 2) are best described by C.C. diagrams such as shown in Fig. 2c. Here, bistability is observed for a single charge state of the defect. For other defects, such as the Fe–Al and C–C pairs in Si (Sect. 1.3), bistability is common to all charge states (Fig. 2d).

It is clear that the existence of the alternate, metastable configurations of these defects can give rise to apparent instabilities in any of their measurable properties, without perfect control of the electronic, thermal and optical history of the samples. Indeed, all of the above-mentioned defects have been discovered through such "experimental metastabilities". This explains why these defects are also often referred to as "metastable defects" in the literature.

1.2. Characteristics and Characterization

The complete characterization of a bistable defect requires the identification of the atomic structures of the various configurations, and of the physical mechanisms for conversion from one configuration to the other. An essential intermediate step is the construction of the C.C. diagram, which does not say much about what the defect is, but which does explain how it works. For the experimentalist, this implies the determination of the electronic (energy level) properties of the defect in its different configurations, and of the kinetics of configurational transformations.

Transformation rates usually display thermally activated behavior:

$$R = R_0 \exp(-E_a/kT). \tag{1}$$

Three controlling physical mechanisms leading to such temperature dependence have been isolated so far:

(i) elementary atomic jump ($R_0 \sim 10^{12}\,s^{-1}$)
(ii) free-carrier capture by multiphonon emission ($R_0 \sim 10^{7}\,s^{-1}$)
(iii) free-carrier emission ($R_0 \sim 10^{13}\,s^{-1}$).

It follows that the preexponential factors of the transformation rates are important parameters of such systems.

A classical tool to isolate and characterize bistable defects is DLTS in conjunction with bias-on/bias-off cooling sequences [5]. Consider a defect with two possible structural configurations, A and B, in both of its charge states, (0) and (+) (Fig. 3). Configuration A is stable when the defect is neutral, whereas B is energetically favorable in the positive charge state. In order to study the energy level properties of a specific configuration (e.g. E_A), the defect is prepared in the proper charge state (e.g. neutral) at around 300 K. This is accomplished by using the voltage applied to the diode to control the free carrier concentration in the active region of the device. The sample is then rapidly cooled to low temperature, and standard DLTS measurements are performed.

One can also use incident light during sample cool down to modulate the Fermi energy in the material and reach the proper charge state. Alternatively, this Fermi level control can be simply achieved by adjust-

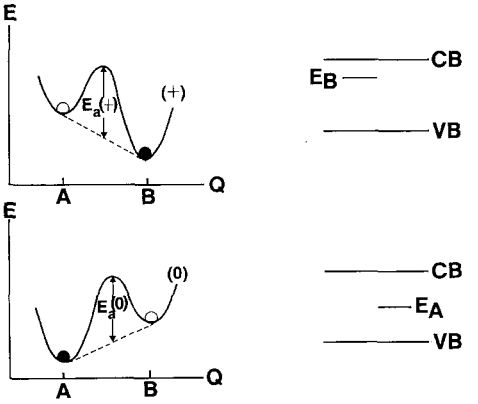

Fig. 3. Schematic drawing of a defect having two charge state controlled configurations

ing the doping of the material. Bulk sensitive techniques, like IR absorption spectroscopy or EPR, can then be used to get insight into the selected atomic defect configuration [3, 6].

Configurational transformations can be studied by freezing the defect in a given charge state and reversing the charge condition at low temperature. The selected configuration becomes metastable, and the system tends to relax towards its new stable structure. One usually performs isochronal anneals in a first stage to isolate the transformation temperature range. The temperature dependence of the transformation rate is then determined by a series of isothermal anneals.

1.3. Physical Models

In most cases of defect bistability reported so far, the defect involved has not been identified microscopically, and the physical origin of the bistability remains unknown. However, as mentioned in the introduction, there are two reported bistable defects in Si for which detailed microscopic models of the two configurations and of the configurational transformations have been proposed. These model defects will be important to understand bistability in other defect systems.

The first of these defects is the interstitial iron-substitutional aluminum pair defect [2]. The C.C. diagram summarizing the bistable electronic properties of this center is shown in Fig. 4, together with the proposed atomic structures. In its stable configuration, the defect has ⟨111⟩ axial symmetry, with Fe sitting in the nearest interstitial site adjacent to Al. The metastable configuration is the ⟨100⟩ oriented structure, where the Fe ion has jumped into the second-nearest-neighbor site. A simple ionic model of the pair, assuming pure electrostatic interaction between point-charge-modeled Fe and Al ions, has been found to provide an adequate fit to most characteristic energies of the system. The bistability observed experimentally is then simply understood as resulting from a change in the relative stability of the nearest- or next-nearest-neighbor positions for the interstitial Fe, caused by a change in Coulomb interaction between Fe^{++} or Fe^{+} interstitial and the negative substitutional acceptor.

The second model bistable defect is the interstitial carbon-substitutional carbon pair [3]. Figure 5 shows the configurational coordinate energy curves and structural models for the two configurations of this defect, as inferred from combined DLTS, EPR, PL, and ODMR experiments. In the A configuration, which is stable in the positive and negative charge states, the defect consists of a substitutional carbon atom next to a carbon interstitially, in which a carbon atom and a silicon atom share a single lattice site. The stable structure of the neutral defect (B configuration) has two substitutional carbon atoms, with a silicon atom squeezed in between, in a position away from the

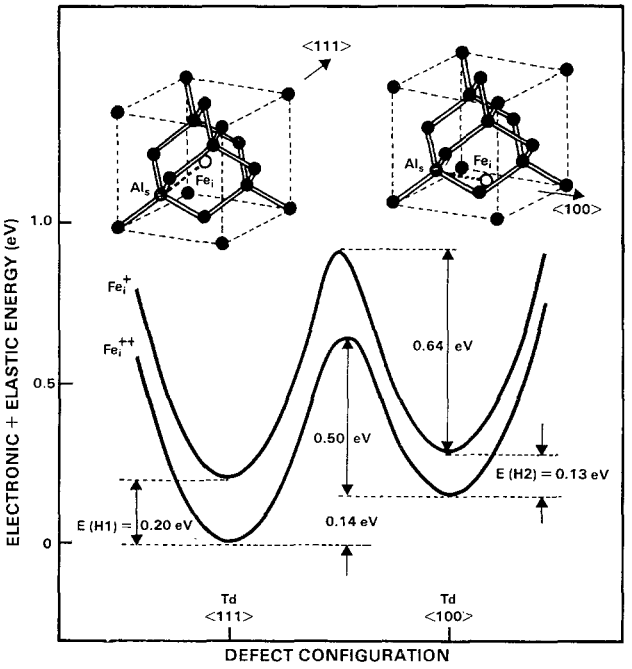

Fig. 4. Configuration coordinate diagram and microscopic model for the Fe–Al pair in Si

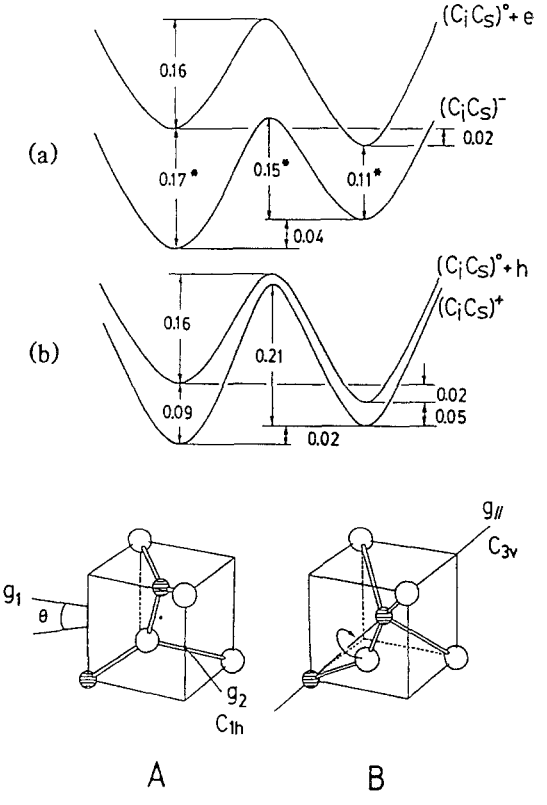

Fig. 5. Models and configurational-coordinate energy curves for the two configurations of the C–C pair in Si, after [3]

bond center. As can be seen in Fig. 5, the conversion from one configuration to the other requires only a simple bond switching Si–Si↔Si–C.

2. Defect Bistability and Thermal Donors in Silicon

After this general introduction to defect bistability, we now turn to a more detailed description of a specific system – bistable thermal donor defects in Si – where our current understanding is far less advanced.

2.1. Background

Thermal donors (TD's) are double donor centers formed upon heating oxygen-rich Si in the 350°–500 °C temperature range. Although their detailed atomic and electronic structure is still unknown, it is generally agreed that TD's are small clusters of atoms formed at the early stages of oxygen precipitation [7]. Up to nine different species (TD1, ..., TD9) have been resolved with IR, with successively decreasing ionization energies vs. heat treatment time [8]. This observation is usually explained in terms of a family of agglomerates, incorporating a growing number of oxygen [9] or silicon [10] atoms.

Recent IR studies [6, 11] have revealed that TD1 (and possibly TD2) display bistable properties. Because TD1 represents the core of the TD family, a detailed study of its bistable behavior can be expected to provide new insight into the structure of TD defects. In the following, we describe the results of our DLTS studies on this new system [12].

2.2. DLTS *Results*

The samples used in this study were prepared from P-doped ($n \sim 2 \times 10^{15}\,\mathrm{cm}^{-3}$) Cz Si. Heat treatments were performed at 450 °C for various durations: 30 min, 1 h, 2 h, 5 h. Such short anneals are expected to generate large relative concentrations of the smaller TD clusters [8].

Figure 6 presents a comparison of two DLTS spectra obtained on a single Au diode of a 30 min annealed sample. Spectrum (a) was recorded after cooling the sample from 350 K to the initial measurement temperature (40 K) with a reverse voltage applied to the Schottky gate. This is a standard procedure for DLTS measurements, which monitor the thermal reemission of trapped carriers in a semiconductor junction depletion region. A single feature labeled $E(0.15)$ is detected, which in the past has been assigned to the second ionization state of the TD [13]. DLTS is unable to resolve the energy levels of individual TD species, so that the $E(0.15)$ signal represents the total response of the whole distribution of cluster sizes.

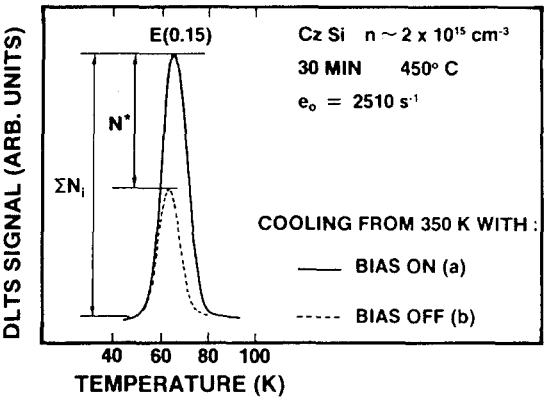

Fig. 6. DLTS spectra for thermal donors in Si, showing the existence of bistable TD species

When the same diode is cooled down with no applied bias, significant changes are observed in the subsequent DLTS run recorded under the same experimental conditions. As shown by spectrum (b) in Fig. 6, the main unexpected effect is a strong reduction (~50% for this sample) of the signal intensity. Moreover, it can be seen that the residual DLTS emission signal is shifted to lower temperatures. The $E(0.07)$ DLTS signal, associated with the first ionization state of the defects, was found to behave similarly.

Results such as shown in Fig. 6 were obtained from all investigated diodes. However, the amplitude of the phenomenon was found to depend critically upon the duration of the 450 °C heat treatment. More precisely, maximum relative change in DLTS signal intensity, $N^*/\sum N_i$ (see Fig. 6 for the notation) was observed for the shortest anneals [12]. As dicussed below, these observations are consistent with the results of IR studies, and provide strong additional evidence for the existence of bistable centers in the early stages of TD formation.

When diodes are cooled to low temperature with an applied reverse bias, no free electrons are present in the active region of the device. All TD species are frozen in the usual double donor configuration. Hence, all TD clusters ($\sum N_i$) contribute to the $E(0.15)$ DLTS signal. Whe the bias is off during sample cool down, bistable TD (BTD) centers (N^*) transform to an other configuration, with different electronic properties (not seen in DLTS). Their contribution to the $E(0.15)$ signal therefore disappears. The observed shift to lower temperatures of the residual DLTS feature indicates that BTD clusters are those having the larger ionization energies, i.e. the smaller sizes. The same conclusion is drawn from the annealing time dependence of $N^*/\sum N_i$, which indicates that the relative concentration of TD clusters is skewed towards larger sizes by increasing the 450 °C heat treatment time.

Configurational intraconversions were then explored in more detail. For these measurements, we used the diode whose characteristic DLTS spectra are shown in Fig. 6. The conversion from the usual double donor configuration to the other one was investigated by studying the transformation from state (a) to state (b) of the system, as defined in Fig. 6. To do so, the sample is first cooled from high temperature (350 K) to a temperature T with applied reverse bias. The diode is then shorted, maintained for a time t at this temperature, and rapidly cooled to 40 K. Configurational changes are then detected as changes in the $E(0.15)$ DLTS peak height. The reverse conversion is studied via transition b→a, i.e. using zero bias cooling and reverse bias anneals.

Isochronal (5 min) anneals were first performed to reveal the transformation temperatures. The results of these experiments are displayed in Fig. 7. It is seen that both transitions occur in a single stage, but not in the same temperature range: 190–240 K for transition a→b, 270–320 K for transition b→a. The anneal time t was then varied at fixed temperatures T (chosen around the transition temperatures) to explore the configurational transformation rates. The isothermal annealing kinetics were found to be first order. Arrhenius plots for both transitions are shown in Fig. 8. The reaction rates are consistent with the following relations:

$$R = 8 \times 10^3 \exp\left[-\frac{0.3\,\text{eV}}{kT}\right] \text{s}^{-1} \quad \text{for} \quad \text{a}\rightarrow\text{b}, \tag{2}$$

$$R = 5 \times 10^{13} \exp\left[-\frac{1\,\text{eV}}{kT}\right] \text{s}^{-1} \quad \text{for} \quad \text{b}\rightarrow\text{a}. \tag{3}$$

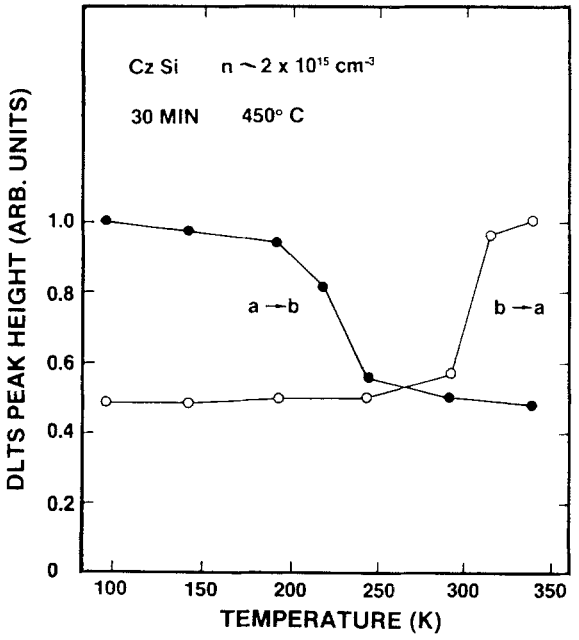

Fig. 7. Isochronal (5 min) annealing data for BTD configurational intraconversions

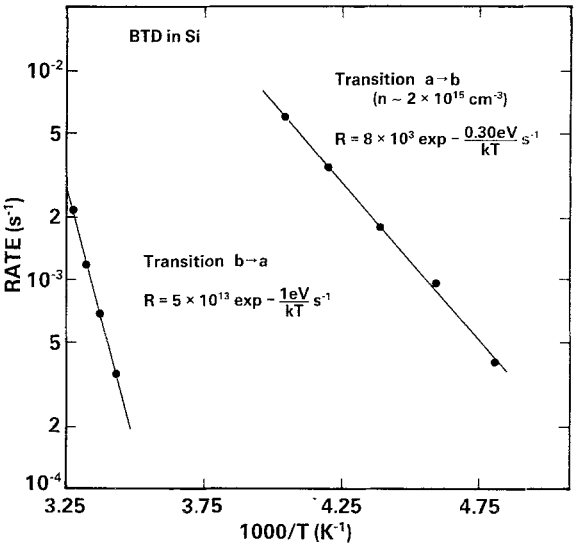

Fig. 8. Arrhenius plots of BTD configurational transformation rates

Data for 5 min isochronal annealing were also collected for a more lightly doped sample ($n \sim 5 \times 10^{14}\,\text{cm}^{-3}$). No difference could be detected for the b→a transformation. However, transition a→b was found to be significantly (~40 K) shifted to higher temperatures.

2.3. *Configuration Coordinate Model*

A schematic C.C. description of BTD centers compatible with the above data is shown in Fig. 9. The A configuration represents the usual, double donor configuration of the defect. The bistable properties originate in the existence of an other, lower energy configuration (B) for the neutral state (BTD^0). The DLTS data shown in Fig. 6 are then readily explained in the following way: When cooling the sample with no

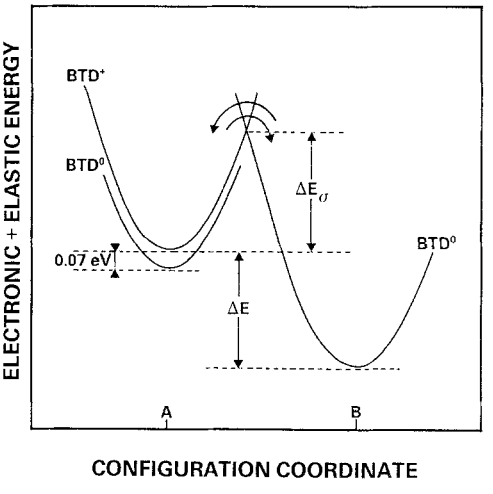

Fig. 9. Schematic configuration coordinate description of BTD electronic properties

voltage on the Schottky gate, free electrons can be captured by BTD^+ at the beginning of the cooling procedure, when the barrier ($\Delta E\sigma$) to the reaction:

$$(BTD^+)_A + e^- \rightarrow (BTD^0)_B \tag{4}$$

can be surmounted. The defect becomes self-trapped into configuration B, which has no energy level in the upper half of the band gap. In the bias on cooling experiment, there are no electrons available for capture in the depleted region of the diode. At low temperature when the bias is removed to allow trap filling, only configuration A of BTD^0 can be accessed:

$$(BTD^+)_A + e^- \rightarrow (BTD^0)_A, \tag{5}$$

and the usual DLTS signals are detected.

Within this model, the rate of the a→b transformation corresponds to the rate of the capture process described by reaction (4). This readily explains the free carrier concentration effect on the transformation temperature range outlined above. The small value of the preexponential factor – $R_0 = 8 \times 10^3\ s^{-1}$ compared to $R_0 \sim 10^7\ s^{-1}$ expected for a free carrier capture process – can be explained by a Fermi level (E_F) effect on the population of the $E(0.15)$ state. Indeed, at the elevated temperatures (210–250 K) of the measurements, E_F lies below the $E(0.15)$ level in these samples. The reaction rate must then be written as:

$$R = \sigma_n n v_{th} \left[1 + \exp\left(\frac{E(0.15) - E_F}{kT} \right) \right]^{-1}, \tag{6}$$

where σ_n is the capture cross section and v_{th} the electron thermal velocity. The last term in (6) represents the fraction of BTD defects in the singly positively charged state. Rewriting $\sigma_n = \sigma_n(\infty) \times \exp\left(- \frac{\Delta E\sigma}{kT} \right)$, and considering that $\exp\left(\frac{E(0.15) - E_F}{kT} \right) \gg 1$, expression (6) becomes:

$$R = \sigma_n(\infty) n v_{th} \frac{n}{N_c} \exp\left(- \frac{\Delta E\sigma - 0.15\,eV}{kT} \right), \tag{7}$$

where N_c stands for the density of states in the conduction band. The $\frac{n}{N_c}$ term in the preexponential factor accounts for the small value obtained experimentally. Relations (2) and (7) can then be used to extract $\Delta E\sigma$. Taking into account the temperature dependences of v_{th} and N_c, one finally obtains $\Delta E\sigma = 0.47\,eV$.

Transformation b→a corresponds to the reverse (electron emission) reaction):

$$(BTD^0)_B \rightarrow (BTD^+)_A + e^-. \tag{8}$$

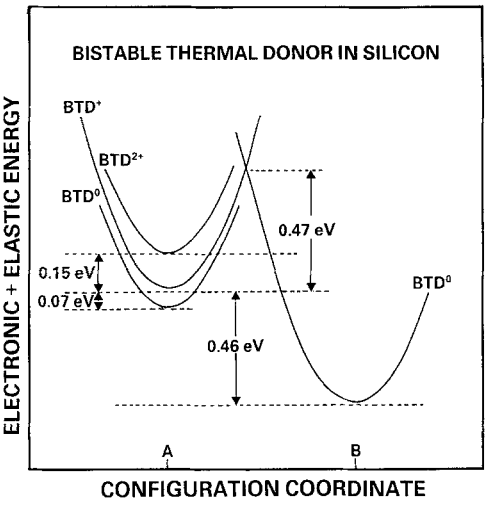

Fig. 10. Configuration coordinate diagram for BTD defects in Si

The rate of this reaction can be written as:

$$R = \sigma_n N_c v_{th} \exp - \frac{\Delta E}{kT}$$

$$= \sigma_n(\infty) N_c v_{th} \exp\left(- \frac{\Delta E + \Delta E\sigma}{kT} \right), \tag{9}$$

where ΔE is the binding energy of the electron trapped in configuration B (see Fig. 9). We note that the preexponential factor of the b→a transformation rate is indeed in the range expected for a free carrier emission process. Comparing (3) and (9), one calculates $\Delta E = 0.46\,eV$, leading to the C.C. model of the defect shown in Fig. 10.

2.4. Discussion

As mentioned previously, there is currently no accepted model for the structure of TD defects. As a result, the bistability displayed by the smaller clusters cannot be understood on a microscopic scale. This is a typical illustration of the more general situation where it is difficult to go beyond a phenomenological C.C. description of the defect bistability.

However, it is interesting to discuss the results of cluster calculations performed on the substitutional oxygen-oxygen pair in Si, which has been considered as a possible core for the TD defects [14]. Those calculations revealed the existence of two possible configurations of the pair, with different electronic properties: one stable configuration, where the two oxygens are sufficiently far apart so as to interact only weakly and introduce no electrical level in the gap, and one metastable configuration, where the oxygens are moved closer together, thereby interacting more strongly, and having shallow double donor characteristics. It was proposed that the presence of surrounding interstitial oxygens could stabilize the interacting O–O

structure, causing the double donor activity of the observed TD's.

The electronic properties of BTD defects closely fit into this picture. We note, however, that other models for the core, e.g. containing more than two oxygen atoms, could possibly give the same results [15]. Considering the fact that the B configuration of BTD's then corresponds to this unperturbed core, it can be anticipated that any direct experimental information from this configuration could provide important clues towards a microscopic understanding of TD's in Si.

3. Concluding Remarks

Bistability is today an important area of research in the field of defects in semiconductors. Many bistable defects have already been reported, both in silicon and compound semiconductors. As more experimentalists have become aware of their existence, and trained in their detection, it can be anticipated that other examples will be discovered in the near future.

The simplest method to isolate such defects is the use of DLTS in conjunction with bias-on/bias-off cooling sequences. The results can be used to construct a configuration coordinate diagram for the defect, which summarizes its electronic properties. However, it is clear that a microscopic description of the various defect configurations, and configurational transformation reactions, cannot be obtained without additional spectroscopic data. The study of bistable thermal donor defects reported in this paper is indeed a good illustration. A remarkable combined DLTS–EPR–PL –ODMR characterization approach has been demonstrated for the bistable carbon-carbon pair defect in silicon. This work will serve as a model example for future research in this area.

References

1. For a recent review of bistable defects in silicon, see A. Chantre: In *Defects in Electronic Materials 1987*, ed. by M. Stavola, S.J. Pearton, G. Davies, MRS Symposia Proceedings (North-Holland, New York 1988) p. 37
2. A. Chantre, D. Bois: Phys. Rev. B **31**, 7979 (1985)
3. L.W. Song, X.D. Zhan, B.W. Benson, G.D. Watkins: Phys. Rev. Lett **60**, 460 (1988)
4. A. Chantre, L.C. Kimerling: Appl. Phys. Lett. **48**, 1000 (1986) L.W. Song, B.W. Benson, G.D. Watkins: Phys. Rev. B **33**, 1452 (1986)
5. J.L. Benton, M. Levinson: In *Defects in Semiconductors II*. ed. by S. Mahajan, J.W. Corbett (North-Holland, New York 1983) p. 95
6. Ya.I. Latushko, L.F. Makarenko, V.P., Markevich, L.I., Murin: Phys. Status Solidi A **93**, K 181 (1986)
7. W. Kaiser, H.L. Frisch, H. Reiss: Phys. Rev. **112**, 1546 (1958)
8. R. Oeder, P. Wagner: In Ref. [5], p. 171
9. A. Bourret: J. Electron. Mater. **14a**, 129 (1984)
10. D. Mathiot: Appl. Phys. Lett. **51**, 904 (1987)
11. B. Pajot, R.C. Newman, P. Wagner: Private communication
12. A preliminary report on this work is given in A. Chantre: Appl. Phys. Lett. **50**, 1500 (1987)
13. L.C. Kimerling, J.L. Benton: Appl. Phys. Lett. **39**, 410 (1981)
14. G.G. De Leo, C.S. Milsted, Jr., J.C. Kralik: Phys. Rev. B **31**, 3588 (1985)
15. G.D. Watkins: Private communication