PHYSICAL REVIEW B VOLUME 41, NUMBER 12 15 APRIL 1990-II

Hysteresis and Franck-Condon relaxation in insulator-semiconductor tunneling

W. B. Fowler, J. K. Rudra,* M. E. Zvanut,† and F. J. Feigl‡

Department of Physics and Sherman Fairchild Laboratory, Lehigh University, Bethlehem, Pennsylvania 18015

(Received 29 December 1989)

We analyze electron tunneling between semiconductor bands and insulator traps at a semiconductor-insulator interface in terms of a Born-Oppenheimer model. For insulator traps which exhibit large electron-phonon interactions this model predicts tunneling processes analogous to optical Franck-Condon transitions, i.e., tunneling followed by atomic relaxation at the defect. Such ideas go back to Gurney's treatment of electrolysis, but have not appeared in the current interface literature. We estimate the relaxation energies for a model of the E' center in silicon dioxide and argue that the hysteresis observed by Zvanut *et al.* in band-to-trap tunneling in Si-SiO_2 most likely arises from such a process, which we call *hysteretic tunneling*. We suggest that such processes should occur in other cases involving insulating defects which exhibit large electron-lattice coupling.

I. INTRODUCTION

The important role of the silicon–silicon-dioxide interface in semiconductor devices has made it an active subject of research for many years. Processes involving electron and hole tunneling are significant in the science and technology of this interface. Of particular interest are "band-to-trap" tunneling processes involving electrons or holes in the silicon bands and defects in the silicon dioxide which are near the interface.

Woods and Williams[1] conducted an important early study in which they caused holes to be trapped in the oxide near the interface and suggested that oxygen vacancies were responsible. Many papers have since been written[2–4] on band-to-trap tunneling, from a number of perspectives.

Meanwhile, extensive research has been carried out on defects in bulk silicon dioxide, and on silicon–silicon-dioxide interface defects. Early electron-paramagnetic-resonance studies in quartz[5,6] were supplemented by studies in the bulk silicon dioxide glass[7] and by theoretical models for oxygen-vacancy defects[8–10] which incorporated large lattice relaxations of oxygen vacancies in different charge states. More recently the arguments of Woods and Williams have been extended.[11] At present the native oxide had not been removed. The data were analyzed in terms of a model in which the traps are located at the interface between native and deposited oxide.

Recently Zvanut and co-workers[12,13] conducted an extensive study and analysis of band-to-trap tunneling in a particular Si-SiO_2 system. The samples were prepared by sputter-depositing silicon dioxide onto silicon from which the native oxide had not been removed. The data were analyzed in terms of a model in which the traps are located at the interface between native and deposited oxide.

One significant feature of this study was a strong hysteresis observed in filling and emptying the traps during an isochronal bias stress cycle. Similar effects have also been observed in p-channel transistors[14] and dual dielectric memory devices.[15] Zvanut and co-workers analyzed their data on the basis of a single defect and a tunneling process which involved subsequent atomic relaxations, similar to Franck-Condon optical transitions at defects in insulators.[16] This process may be called *hysteretic tunneling*. The purpose of this paper is to present the theoretical arguments concerning hysteretic tunneling in some detail, and to relate these arguments to the level positions determined theoretically and experimentally in band-to-trap tunneling.

II. HYSTERETIC TUNNELING

The notion of hysteretic tunneling associated with the Franck-Condon principle goes back to a classic paper by Gurney[17] on the quantum mechanics of electrolysis. One ingredient of this process is the neutralization of positive or negative ions by electron tunneling. In this paper Gurney argued that for any hydrated positive ion the neutralization potential is less than the metal work function, while for a negative ion the neutralization potential is greater than the work function, owing to the Franck-Condon principle in the tunneling process.

Later Silsbee put these ideas on a firmer footing in an analysis of tunneling of molecular defects in insulating crystals.[18] Silsbee argued that the physics of tunneling in strongly relaxed systems follows exactly the physics of optical transitions in such systems.

Other papers containing similar notions include the treatment of oxidation-reduction reactions involving electron transfer by Marcus,[19] a treatment of the photoemission of electrons from tellurium into sulfur by Williams and Sanchez Sinencio[19] and an analysis of inverse-photoemission spectroscopy at a metal-electrolyte interface by McIntyre and Sass.[20]

It should be noted here that tunneling processes are often thought to take place slowly, whereas optical transitions are argued to occur rapidly compared with lattice relaxation times. Therefore one might be tempted to argue that a "fast" optical process may be followed by defect relaxation, while "slow" tunneling may occur only

 ©1990 The American Physical Society

between two relaxed states (a zero-phonon process). Notwithstanding that tunneling time is an ill-defined quantity[21] which continues to be subject to controversy, it is important to note that Gurney-Silsbee-type arguments do not imply that the tunneling takes place either rapidly or slowly. The same is true of optical processes; for example, the tunneling-recombination luminescence between distant defects in alkali halides can have a radiative lifetime of *hours,* while still obeying the Franck-Condon principle.[22]

The initial and final states involved in tunneling must be carefully considered. For example, we treat a tunneling process in which an electron from the valence band of silicon tunnels to a positively charged silicon dioxide defect, thereby neutralizing the defect. The initial state of the *system,* $|I\rangle$, is the combination of a filled valence band of silicon, with appropriate atomic positions, and a positively charged oxide defect, again with appropriate atomic positions. This initial state is thermodynamically well defined in that the atoms are fully relaxed. The state of the system after tunneling, which we call the final state $|F\rangle$, is the combination of a hole in the silicon valence band and a neutral oxide defect, with *some* atomic positions. $|I\rangle$ and $|F\rangle$ may be written in Born-Oppenheimer form as

$$|I\rangle=|\phi\rangle_{\mathrm{Si},i}\,|X\rangle_{\mathrm{Si},i}\,|\phi\rangle_{d,i}\,|X\rangle_{d,i}\ , \tag{1}$$

$$|F\rangle=|\phi\rangle_{\mathrm{Si},f}\,|X\rangle_{\mathrm{Si},f}\,|\phi\rangle_{d,f}\,|X\rangle_{d,f}\ . \tag{2}$$

Here $|\phi\rangle$ is an electronic wave function, $|X\rangle$ a vibrational wave function. i stands for the initial state, f for final, and d for defect.

The tunneling process is energy conserving and the tunneling operator acts only on the electronic states. Thus the tunneling matrix element between $|I\rangle$ and $|F\rangle$ involves an electronic matrix element containing the tunneling operator, times the overlap of vibrational wave functions. It is likely that there will be little atomic relaxation in the silicon associated with the removal of an electron from the valence band, so we may assume that the vibrational overlap between $|X\rangle_{\mathrm{Si},i}$ and $|X\rangle_{\mathrm{Si},f}$ will be nearly unity. We then concentrate on the vibrational overlap between $|X\rangle_{d,i}$ and $|X\rangle_{d,f}$. For vibrational wave functions associated with atomic coordinates which are nearly the same in initial and final states, this overlap is large. When the atomic coordinates are very different, this overlap is small. The former case represents a transition with no change in configuration (the Franck-Condon transition), while the latter case represents a transition between fully relaxed states (the zero-phonon transition). In the strong-coupling case the Franck-Condon transition is much stronger than the zero-phonon transition.

Such effects can be seen readily in a simple one-dimensional linear coupling model.[23] In this model the strength of the electron-phonon coupling at the defect is measured by the Huang-Rhys[24] factor S. S is the number of phonons (of energy $\hbar\omega$) emitted after the Franck-Condon transition takes place, so $S\hbar\omega$ is the energy difference between the zero-phonon line and the Franck-Condon peak. We first suppose that the defect level is deep within the valence band, that the density of valence-band states is constant, and that the electronic tunneling matrix element is energy independent. We also assume that $T=0$ K, so that there are no phonons associated with the initial state. Then, in this model, the tunneling probability to a state with p phonons excited is proportional to [23] W_p, defined by

$$W_p=(S^pe^{-S})/p!\ . \tag{3}$$

According to the Franck-Condon principle, the maximum probability is associated with $p\approx S$. The ratio of the strength of the zero-phonon transition to that of the $p=S$ transition is then

$$W_0/W_S=S!/S^S\ . \tag{4}$$

If we consider a case with $S=20$ as a representative value for a strongly coupled system, then the ratio in Eq. (4) equals 2.3×10^{-8}. On the other hand, for a weakly coupled system (e.g., $S=3$) this ratio equals 2.2×10^{-1}.

We now consider a strongly coupled oxide defect whose unrelaxed (Franck-Condon) energy lies within the semiconductor band gap but whose relaxed (zero-phonon) energy lies within the semiconductor valence band. In the absence of an electrical bias, Franck-Condon tunneling will not occur since energy would not be conserved, while energy-conserving tunneling has a low probability because of the small vibrational overlap. If an appropriate bias is applied which brings the Franck-Condon state to an energy within the semiconductor valence band, then the higher-probability tunneling process can occur.

To determine the feasibility of this model, we must determine the relative probability of the zero-phonon process versus the Franck-Condon process. Equation (4) provides the answer if the coupling strength S is known. S may be estimated from the relaxation energy as the ratio of relaxation energy to a typical phonon energy. Our calculations of the relaxation energy associated with a particular defect in the SiO_2/Si system are discussed in the following section.

III. ELECTRICAL LEVELS IN HYSTERETIC TUNNELING

The *defect level* is a measure of the energy involved in changing the charge state of a defect when both initial and final states of the system are in thermodynamic equilibrium. Thus the energy denoted by $E(n/(n+1))$ is defined with respect to the vacuum by

$$E(n/(n+1))=E_n-E_{n+1}\ , \tag{5}$$

where E_m is the total energy of the relaxed system when its charge state is m. $E(n/(n+1))$ may also be defined with respect to the conduction or valence band edge by including that energy in Eq. (5).

The occupancy of the defect level in thermodynamic equilibrium is determined by the chemical potential, or Fermi energy. If the Fermi energy of the system is below $E(n/(n+1))$ but above the next lower level, the charge state of the defect will be $(n+1)$; similarly, if the Fermi energy is above $E(n/(n+1))$ but below the next higher level, the charge state of the defect will be n.

As discussed in Sec. II, the tunneling process may not take the defect directly into the equilibrium state, but into an unrelaxed state after which vibrational relaxation occurs. In order to describe both relaxed and unrelaxed states, we introduce a new notation. We define $E_{n(m)}$ as the energy of the system in the charge state n and the equilibrium configuration corresponding to the charge state m. Then, $E_{n(n)}$ corresponds to the previously defined E_n. When the system is in the unrelaxed state, m differs from n by ± 1.

We now apply these ideas to the model which has been suggested to explain the observed hysteresis in the Si-SiO_2 system[13] mentioned earlier. The defect involved is taken to be a surface-like or "hemi" E' center consisting of a dangling bond on the threefold coordinated Si in silicon dioxide. This defect is paramagnetic and neutral when one electron occupies the dangling bond, and it is diamagnetic and positively charged $(+1)$ when the dangling bond is empty. It has been suggested by Zvanut *et al.*[13] that this defect exchanges charge with the valence band of silicon.

To analyze this defect we consider the states $n(m)$, using the notation defined above. Then 0 (0) refers to the neutral E' center in its minimum energy configuration. Similarly, $0(+)$ refers to the neutral E' center in the minimum-energy configuration of the positive charge state; $+(0)$ refers to the positively charged defect in the minimum-energy configuration of the neutral charge state; and $+(+)$ refers to the positively charged defect in its minimum-energy configuration. The corresponding energies and their locations on the simple configuration coordinate diagram of Fig. 1 are $E_{0(0)}$ at A, $E_{+(0)}$ at B, $E_{+(+)}$ at C, and $E_{0(+)}$ at D. The left-hand well represents the total energy of the system consisting of a neutral oxide defect and a hole in the Si valence band. The right-hand well represents the energy for the positively charged oxide defect with an electron in the Si valence band. In this example, the two wells are chosen to be in alignment under zero applied bias. Under this circumstance, tunneling between zero-point vibrational states conserves energy. As shown in Figs. 1(a)–1(c), states A and B and states C and D may be brought into communication with each other by an applied bias V_a. The charging-discharging process which results is described as follows.

During discharging, a bias is applied such that the unpaired electron of the defect tunnels to the valence band of Si. This is represented in Fig. 1(b) where the defect transformation from state A [0(0)] to state B [$+(0)$] is illustrated. Subsequent atomic relaxation reduces the energy and the defect becomes $+(+)$ at C. During charging, a bias is applied such that an electron from the valence band of Si tunnels to the defect. This is shown in Fig. 1(c) where the defect transformation from state C to D is illustrated. The defect then returns to 0(0), at A, after reducing its energy by atomic relaxation.

The situation illustrated in Figs. 1(a)–1(d) represents the most probable tunneling event. The vibrational level at which tunneling takes place is the one in which the vibrational wave-function overlap is maximum, i.e., the Franck-Condon process. Transitions between other vibrational levels are possible with a lower probability, as described in Sec. II.

Returning to the general case, we redefine the level position with respect to the vacuum given originally in Eq. (5) as follows:

$$E(n(m)/n'(m'))=E_{n(m)}-E_{n'(m')} . \tag{6}$$

Our case, however, addresses the intermediate levels

$$E(0(0)/+(0))=E_{0(0)}-E_{+(0)} \tag{7}$$

and

$$E(0(+)/+(+))=E_{0(+)}-E_{+(+)} . \tag{8}$$

Thus, when the $0(0)/+(0)$ level is moved above the Fermi level the neutral defect is converted by tunneling to a positive one, and when the $0(+)/+(+)$ level is moved below the Fermi level the positive defect is con-

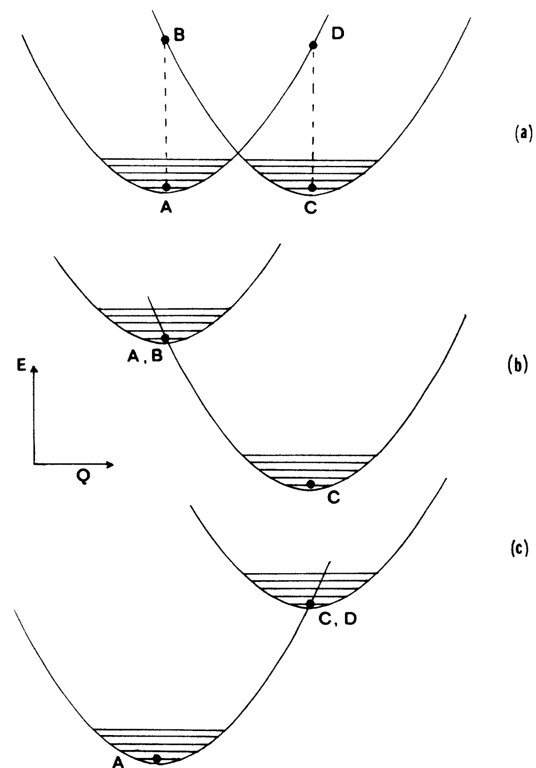

FIG. 1. Schematic configuration coordination diagrams for a strongly coupled defect near a semiconductor interface under the special circumstance that with zero bias, tunneling between zero-point vibrational states conserves energy. In each case the left-hand well represents the total energy of the system when the defect has a positive charge. The first few vibrational levels are shown. In terms of the energies defined in the text, A corresponds to $E_{0(0)}$, B to $E_{+(0)}$, C to $E_{+(+)}$, and D to $E_{0(+)}$. Panel (a) represents the situation of zero bias V_a, (b) a bias leading to electron tunneling from the defect to the silicon, and (c) a bias leading to electron tunneling from silicon to the defect.

verted to a neutral one. It is the energy difference between these two levels, $[E_{0(+)}-E_{+(+)}]-[E_{0(0)}-E_{+(0)}]$, which accounts for the hysteretic tunneling of this model of the process.

We have computed these energies as well as the thermodynamic level $[0(0)/+(+)]$ for a $Si_4O_{12}H_9$ cluster chosen to simulate the hemi E' center, using the semiempirical quantum-mechanical technique[25] MINDO/3 and its open-shell version MOPN. The results of these calculations are shown in Fig. 2, where the numbers are computed relative to the vacuum energy and the experimental valence and conduction band edges are shown by dotted lines.

The calculation shows that for a particular defect, the hemi E' center, the relaxation energy is ≈ 0.5 eV. If a typical phonon energy of ≈ 0.02 eV is considered, a Huang-Rhys factor, S, equal to 25 results, representative of a strongly coupled system for which the ideas based on Franck-Condon tunneling developed above are expected to apply. It is reasonable, then, to interpret the trapping behavior observed by Zvanut *et al.* as hysteretic tunneling.

A further important point should be made: even if the (energy-conserving) zero-phonon tunneling process is improbable, it will eventually occur in the absence of competing processes if one waits long enough. The zero-phonon tunneling rate depends on the size of the electronic tunneling matrix element. We again use the numerical illustration of Sec. II, with $S=20$, as an example. If the tunneling rate for the Franck-Condon process is 1 sec^{-1}, then the corresponding rate for the zero-phonon process will be $\approx 10^{-8}$ sec−1 and this process would take place on the order of once per year. If on the other hand the Franck-Condon rate is, say, 10^6 sec^{-1}, then the zero-phonon rate will be $\approx 10^{-2}$ sec^{-1}, or once per two minutes. Thus the measurement time becomes in principle an important experimental parameter. A sufficiently short measurement time will eliminate the non-Franck-Condon processes, whereas a long measurement time will allow such processes to occur. Then the observed energy distribution of the defects will broaden, reflecting the occurrence of both types of processes.

Furthermore, if the defects are distributed in distance from the interface, one might find that those closest to the interface will recombine in the absence of bias through a zero (or few) phonon process, while those farther from the defect will remain "stable" until a bias is applied. Were one to measure the recombination as a function of time in the absence of bias, one should see a nonexponential decay similar to that observed in tunneling recombination of alkali halides.

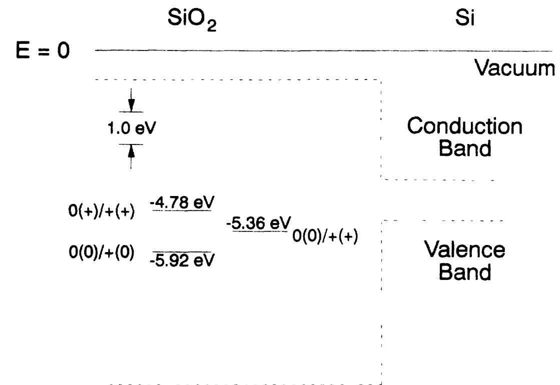

FIG. 2. Level positions computed for a hemi E' center as described in the text. Shown in dotted lines are the experimental conduction and valence-band edges for silicon and for silicon dioxide.

IV. CONCLUSIONS

We have argued that the hysteresis observed in band-to-trap tunneling in a silicon–silicon-dioxide system by Zvanut and co-workers[12,13] can be explained by a Franck-Condon process in which electron tunneling is followed by atomic relaxation at the defect. We expect this description of band-to-trap tunneling to be valid in other cases which involve insulator defects with large electron-lattice coupling. Thus, for example, the origin of hysteresis observed in *p*-channel transistors[14] and dual dielectric memory devices[15] deserves further consideration.

ACKNOWLEDGMENTS

The authors acknowledge helpful discussions with Arthur H. Edwards, Frank Ham, and George D. Watkins and comments by R. Williams and E. H. Poindexter. This research was supported in part by the Electronic and Solid State Science Program of the U.S. Office of Naval Research. M.E.Z. gratefully acknowledges financial support from the IBM Corporation.

*Present address: Department of Physics/Engineering, Xavier University of Louisiana, 7325 Palmetto Street, P.O. Box 82C, New Orleans, LA 70125.

†Present address: Department of Physics and Astronomy, University of North Carolina, Phillips Hall, Chapel Hill, NC 27599-3255.

‡Deceased.

[1]M. H. Woods and R. Williams, J. Appl. Phys. **47**, 1082 (1976).

[2]S. Manzini and A. Modelli, in *Insulating Films on Semiconductors,* edited by J. F. Verweij and D. R. Wolters (North-Holland, Amsterdam, 1983), p. 112.

[3]I. Lundstrom and C. Svensson, J. Appl. Phys. **43**, 5045 (1972).

[4]H. Koster, Jr. and M. Schmidt, Mater. Sci. Forum **38-41**, 1331 (1989).

[5]R. A. Weeks and C. M. Nelson, J. Am. Ceram. Soc. **43**, 399 (1960); C. M. Nelson and R. A. Weeks, *ibid.* **43**, 396 (1960).

[6]R. H. Silsbee, J. Appl. Phys. **32**, 1456 (1961).

[7]D. L. Griscom, Phys. Rev. B **22**, 4192 (1980).

[8]F. J. Feigl, W. B. Fowler, and K. L. Yip, Solid State Commun. **14**, 225 (1974).

[9]K. L. Yip and W. B. Fowler, Phys. Rev. B **11**, 2327 (1975).

[10]J. K. Rudra and W. B. Fowler, Phys. Rev. B **35**, 8223 (1987).

[11]H. S. Witham and P. M. Lenahan, Appl. Phys. Lett. **51**, 1007 (1987).

[12]M. E. Zvanut, F. J. Feigl, and J. D. Zook, J. Appl. Phys. **64**, 2221 (1988).
[13]M. E. Zvanut, F. J. Feigl, W. B. Fowler, J. K. Rudra, P. J. Caplan, E. H. Poindexter, and J. D. Zook, Appl. Phys. Lett. **54**, 2118 (1989).
[14]J. R. Davies, *Instabilities in MOS Devices* (Gordon and Breach, New York, 1981), p. 93.
[15]J. T. Wallmark and J. H. Scott, RCA Rev. **30**, 335 (1969); E. C. Ross and J. T. Wallmark, *ibid.* **30**, 366 (1969).
[16]See, e.g., W. B. Fowler, in *Physics of Color Centers,* edited by W. B. Fowler (Academic, New York, 1968), p. 54.
[17]R. W. Gurney, Proc. R. Soc. London Ser. A **134**, 137 (1931).
[18]R. H. Silsbee, J. Phys. Chem. Solids **28**, 2525 (1967).
[19]R. A. Marcus, J. Chem. Phys. **24**, 966 (1956); R. Williams and F. Sanchez Sinencio, J. Chem. Phys. **52**, 3216 (1970).
[20]R. McIntyre and J. K. Sass, Phys. Rev. Lett. **56**, 651 (1986).
[21]K. K. Thornber, Thomas C. McGill, and C. A. Mead, J. Appl. Phys. **38**, 2384 (1967).
[22]C. J. Delbecq, Y. Toyozawa, and P. H. Yuster, Phys. Rev. B **9**, 4497 (1974); C. J. Delbecq, D. L. Dexter, and P. H. Yuster, *ibid.* **17**, 4765 (1978).
[23]D. B. Fitchen, in *Physics of Color Centers* (Ref. 16), p. 293.
[24]K. Huang and A. Rhys, Proc. R. Soc. London Ser. A **204**, 406 (1950).
[25]R. C. Bingham, M. J. S. Dewar, and D. H. Lo, J. Am. Chem. Soc. **97**, 1285 (1975); P. Bischof, *ibid.* **98**, 6844 (1976); Arthur H. Edwards and W. Beall Fowler, J. Phys. Chem. Solids **46**, 841 (1985).