Hot-Electron Degradation of Bipolar Transistors

by

Noah Zamdmer

Submitted to the
Department of Electrical Engineering and Computer Science
and the Department of Physics
in partial fulfillment of the requirements
for the degrees of

Master of Science in Electrical Engineering and Computer Science,
Bachelor of Science in Electrical Engineering and Computer Science
and
Bachelor of Science in Physics

at the
Massachusetts Institute of Technology
May, 1994

Signature of Author _____
May 6, 1994

Certified by _____
Professor James Chung, Thesis Supervisor

Certified by _____
Dr. Terence Hook, Company Supervisor (IBM Microelectronics)

Accepted by _____
Aaron Bernstein, Thesis Coordinator, Department of Physics

Accepted by _____
F. R. Morgenthaler, Chairman, E.E.C.S. Dept. Committee on Graduate Students

Hot-Electron Degradation of Bipolar Transistors

by

Noah Zamdmer

Submitted to the
Department of Electrical Engineering and Computer Science
and the Department of Physics on September 19, 1994
in partial fulfillment of the requirements for the degrees of
Master of Science in Electrical Engineering and Computer Science,
Bachelor of Science in Electrical Engineering and Computer Science
and Bachelor of Science in Physics.

Abstract

Hot-electron degradation of a bipolar transistor occurs when the transistor's emitter-base junction is reverse biased. Hot carriers generated by high electric fields in this junction create interface states and other defects that degrade a transistor's forward-bias current gain. This thesis aims to give the most precise microscopic description to date of reverse-bias stress of bipolar transistors. It serves as a guide to establishing the reliability of any BJT or BiCMOS technology. The thesis includes a prediction of the rate of emission of hot, damage-producing carriers from the experimental bipolar transistor's emitter-base junction during reverse-bias stress. It is predicted that hot electrons create most of the damage. Experimental data are presented and analyzed and the three types of defects responsible for all changes in the experimental transistor's I-V and C-V curves are identified. Those defects are interface states, positively charged "slow states" and trapped electrons. The same defects are created during the hot-electron injection of MOS capacitors, by the same fluence of hot electrons predicted in this thesis. The thesis includes descriptions of how the above defects accumulate as stress progresses, and how interface states are annihilated after stress is completed. The various ways in which the defects alter the experimental transistor's I-V and C-V curves are explained.

Thesis Supervisor: James Chung
Title: Assistant Professor of Electrical Engineering

Acknowledgments

My experience as a co-op student at IBM was enlightening, fruitful and delightful. My colleagues helped me enthusiastically at every step along the way to this thesis: they introduced me to the laboratory when I was a novice and helped me to arrive at the arguments that are presented in this work. I would especially like to thank Terence Hook for his guidance and encouragement, and for answering numerous technical questions during my stay at IBM. His mentorship was key to my progress as a co-op student. I would like to thank Harvey Allard Jr. and Donald Cook for teaching me how to design experiments in their laboratory, and Steve Furkay and Jeff Johnson for helping me to use computer simulation. I would also like to thank Peter Cottrell and Micheal Schwartz, my managers at IBM, for their assistance and support.

I thank Prof. James Chung, my MIT thesis advisor, for advising me on how to advance my research and publish my results.

To my parents, my tinder

Contents

List of Tables

Chapter 1

Introduction

Reverse-bias stress of the emitter-base junction of a bipolar transistor causes gain degradation at low forward bias. In many digital and analog circuits, and especially in BiCMOS digital circuits, emitter-base junctions are subjected to both momentary and steady-state reverse bias. A modern bipolar transistor with a shallow emitter and a highly doped, low resistance base has high electric fields in its emitter-base depletion region; in reverse bias, those fields are at their highest. Substantial leakage current then flows, and the carriers that make up the current are accelerated to high kinetic energies by the high electric fields. These hot carriers create recombination centers on the Si-SiO_2 interface where the metallurgical junction meets the passivation oxide along the emitter perimeter. The recombination centers add a recombination current component to I_b, the base current, without affecting I_c, the collector current. Thus as I_b rises, gain decreases. As technology improves and the dimensions of all semiconductor devices decrease, the electric fields in the emitter-base junction of bipolar transistors rise and degradation by hot carriers becomes more and more severe.

Stress of bipolar transistors causes more than the creation of recombination centers. I_b decreases following stress if the emitter-base junction is at zero bias and decreases dramatically following stress if the junction is forward biased. Researchers have proposed the existence of quickly annealing recombination centers [1], donor-type electron traps [2] and trapped electrons in the passivation oxide [3] to explain the changes in I_b, yet these claims have not been proven. The creation of defects by hot carriers in MOS capacitors has been well studied, but defect creation in MOS capacitors has not been compared with defect creation in BJTs. A precise model of BJT stress, one that would give a reasonable physical explanation for the degradation and relaxation of I_b, has not been proposed.

This thesis aims to present the most precise and detailed micrscopic description to date of reverse-bias stress of the emitter-base junction of bipolar transistors. In Chap. 2 we describe hot-carrier phenomena in general and show why bipolar transistors are always vulnerable to hot-carrier degradation. We summarize the history of reverse-bias stress experiments. In Chap. 3 we predict that hot electrons are solely responsible for defect creation in the experimental bipolar transistor. Hole injection is negligible during the MOSFET hot-carrier stress most similar to BJT reverse-bias stress. Chap.

4 contains a calculation of the electron injection ratio, the ratio of the current of electrons injected into the passivation oxide above the emitter-base depletion region to the reverse current that flows simultaneously through the emitter-base diode during reverse-bias stress. The calculation is based on theories that have explained channel hot-electron injection in MOSFETs. Although it is impossible to measure a current of injected hot electrons, the defects we observe are known to be created by electrons alone, by quantities of hot electrons comparable to those supplied by the predicted injection ratio. Thus we indirectly verify the existing carrier heating and injection theories.

In Chap. 5 we present the transistor parameter changes we observed during reverse-bias stress experiments. We identify the defects that cause the parameter changes and explain how the defects cause them. We explain how the transient behavior of defects following stress causes further changes. With computer simulation, we match simulated transistors with defects to actual, stressed transistors by duplicating the experimental data. In this way, we find the populations and positions of the defects in damaged devices. To prove the physical validity of our conclusions, we show that the same quantities of the same defects are created in MOS capacitors under similar stress conditions.

Chapters 3, 4 and 5 give a nearly complete discussion of carrier heating and defect creation in the experimental bipolar transistor. An examination of the effect of the emitter-base junction electric fields on the recombination current that flows at low forward bias is reserved for Chap. 6. We present experiments that reveal the annealability of those recombination centers in Chap. 7.

We have some novel means at our disposal to draw the conclusions presented in this thesis. Our experimental transistor (see appendix A) has a very long emitter perimeter of five centimeters. Since carriers are heated at the emitter perimeter, the effects of hot-carrier damage are magnified and observable. Not only is the capacitance of the emitter-base junction readily measurable, but the changes in that capacitance induced by hot-carrier stress are also measurable. Thus we increase the amount of information obtainable from the base and emitter terminals alone. Previous experimenters have observed the trapping of charge in the passivation oxide above the emitter-base junction; we use the junction capacitance at zero applied bias to monitor how the shape of the emitter-base space charge region is distorted by that charge. We use the process simulator FEDSS [4] to construct a mathematical replica of the experimental transistor, and the semiconductor equation solver FIELDAY [5] to simulate the effect of charged defects in the passivation oxide on various device parameters (a comparison of the true and simulated experimental transistors is in appendix A).

Chapter 2

Background

Hot-carrier phenomena in solid state devices take place when free electrons and holes are accelerated by high electric fields to kinetic energies many times greater than the thermal energy. High-energy carriers damage a device by activating chemical reactions, displacing atoms in the device's crystal lattice, being trapped by defects and impurities, and creating traps and recombination centers.

Silicon dioxide and the SiO_2-Si interface are particularly vulnerable to hot carriers. Free carriers heated by high electric fields in an oxide film create numerous defects in the film. When carriers are heated in silicon which is in the vicinity of SiO_2, much more damage is created by the relatively few hot carriers that overcome the SiO_2 energy barrier and enter the SiO_2 than the carriers that remain in the silicon. In fact, one can assume that all the damage is created within the film or along its boundary.

A simple example of carrier heating followed by injection into SiO_2 is the avalanche injection of electrons from the p-type substrate of a MOS capacitor into the gate oxide. The sudden application of a large positive voltage to the capacitor gate causes high electric fields in the deeply depleted substrate which accelerate thermally-generated free electrons directly towards the gate. The electrons with the highest kinetic energy surmount the gate oxide energy barrier and produce defects in the oxide (a band diagram is shown in Fig. 1). Other examples of carrier heating and injection into SiO_2 are complicated by the fact that the electric fields that heat electrons and holes don't direct either carrier towards the vulnerable oxide; collisions are necessary to divert hot-carriers from the heating region to the SiO_2. In spite of this, all the observed hot-carrier damage is located in the oxide. The close proximity of a high-field region and silicon dioxide is the necessary condition for the occurrence of significant hot-carrier damage in a solid state device.

Virtually all bipolar transistors, both ancient and modern, fulfill that condition. An SiO_2 film is almost always used to passivate the silicon wafer surface at the emitter-base metallurgical junction. The electric fields in the emitter-base depletion region near the passivation oxide tend to be high. The emitter-base junctions of some of the bipolar transistors on which hot-electron experiments have been performed are shown in Fig. 2. They are transistors of varying degrees of sophistication, with self-aligned emitters (diodes D and E) and non-self-aligned emitters. They are arranged

in order of decreasing extrinsic base[1] to emitter distance, since the trend in bipolar modernization is the decrease of this distance.

In diodes C, D and E, the extrinsic base forms a heavily doped p-n junction with the emitter adjacent to passivation oxide along the emitter perimeter. As one can imagine, the electric fields in that junction are very strong; even with a low reverse bias applied to the junction, lower than the breakdown voltage, hot electron damage is created.

Diodes A and B have no extrinsic base at all. The p+ regions in diodes A and B are base contacts. Though diodes A and B do not have highly doped extrinsic bases, in both diodes high electric fields exist near silicon dioxide. Diode A, made in the early 1970's, has a diffused base. The peak boron concentration and peak electric fields are at the wafer surface, immediately beneath the passivation oxide film. Diode B, the experimental diode of these authors, has an intrinsic base implanted through passivation oxide at low energy. A low implant energy was chosen to make the base width as narrow as possible; the boron ions barely traverse the oxide during implantation. The peak boron concentration and peak emitter-base junction electric fields again abut SiO_2 at the base surface, where they can produce severe hot-electron damage.

Collins [11] recognized in the late 1960's that the degradation of bipolar transistors induced by reverse-bias stress is caused by hot carriers rather than ion migration or any other phenomenon. His deduction is based on his observations that the damage is proportional to the emitter perimeter, is aggravated by high base doping and is nearly independent of temperature. He also observed that avalanche is neither a necessary nor sufficient condition for the degradation.

The above paragraphs give a very short summary of hot-electron phenomena and describe why bipolar transistors suffer hot-electron damage. That brief discussion must be followed by a detailed model of the hot-carrier degradation of BJTs, which is readily constructible. Defect creation by a hot carrier in a bipolar transistor is a sequence of distinct events: carrier heating in the emitter-base depletion region, injection over an energy barrier into the nearby passivation oxide and reaction with one of a variety of chemical species in the oxide to form a defect. The events are independent of each other. Heating of electrons and holes in the emitter-base depletion region depends solely on the current and electric field distributions in that region at a given bias condition. The height of the passivation oxide energy barrier is modulated only by the electric field in the oxide. Defect creation depends on the kinetic energy of carriers in the passivation oxide, which is determined by the electric field in the oxide and not by the kinetic energy of the carriers prior to injection. That is because the transport of carriers in oxides thicker than 100 Å is controlled by scattering and is not ballistic [12]. The high-field region in the silicon serves as a reservoir from which damage-producing hot carriers are injected into the silicon dioxide, limited by the oxide energy barrier.

The events in the defect creation sequence have been thoroughly investigated and

[1] an extrinsic base is a highly doped region between the base contact and the emitter that reduces the base resistance

are well understood. The carrier heating and carrier injection into SiO_2 that take place during channel hot-electron stress of MOSFETs have been modeled successfully; MOS capacitors have proven to be the ideal vehicle for the study of defect formation in oxide films due to their one-dimensionality. The same events in bipolar transistors have never been studied in the same detail, thus the physics of hot-carrier degradation in bipolar transistors is quite primitive in comparison with the understanding of hot-carrier effects in MOSFETs and MOS capacitors. The wealth of information from MOSFET and MOS studies should be used to explain the results of bipolar experiments, as will be done in this thesis.

The hot-carrier physics of bipolar transistors lags behind that of other devices because modern BJTs lack a gate over the emitter-base junction. In modern transistors (diodes B, C, D and E of Fig. 2), the emitter is doped with a high concentration of arsenic that barely diffuses from the polysilicon emitter contact. The emitter contact extends over the entire emitter-base metallurgical junction, leaving no space for a gate. The lack of a gate hampers the study of hot-carrier effects in BJTs in three ways. First, hot-carrier phenomena are controlled by only the base and emitter terminals, limiting the variety of experiments that can be performed upon BJTs. Second, only probes of those two terminals reveal information about the damage induced by hot-carriers, making it very difficult to distinguish the various possible defects or measure their populations (the collector is distant from the region of carrier heating, uninvolved in defect creation and unaffected by defects). Third, no terminal is available to collect the hot-carriers injected from the emitter-base junction into the overlying oxide. The MOSFET gate collects carriers injected into the gate oxide, from either the drain-substrate junction during channel hot-electron (CHE) injection or the substrate during substrate hot-electron (SHE) injection, allowing the population of hot carriers in a MOSFET to be counted and predictions of that population to be directly verified. Such verification is impossible with ungated BJTs.

Planar transistors of the early 1970's (see diode A of Fig. 2) were large enough to allow a gate to be placed on the passivation oxide over the emitter-base junction. Models of carrier heating and carrier injection into the passivation oxide during the avalanche reverse-bias stress of that junction were verified by collecting the hot carriers that reached the gate. Those experiments, such as the one by Bulucea [13], are similar to the channel hot-carrier stresses of MOSFETs during which the gate current is also measured. Other scientists performed reverse-bias stresses at different gate voltages to vary the electric fields at the emitter-base junction during injection. The work of MacDonald [14] is typical of such experiments. He observed the creation of recombination centers at the silicon-passivation oxide interface at negative, positive and zero applied gate voltages; the trapping of electrons in the passivation oxide at positive gate voltages and the trapping of holes at negative gate voltages. He presented no detailed model of defect creation.

Since the 1970's, bipolar hot-carrier experiments have mainly been performed on the actual, ungated transistors used in VLSI circuits. Due to the lack of a gate, the focus of recent experiments has not been on modeling the events in the defect creation sequence listed above. Relationships between stress time and stress reverse current and the change in device parameters they effect have been formulated by making

assumptions about the types of defects created and the way those defects influence a BJT.

The most dramatic change in device characteristics caused by hot-electron stress is an increase in base current, I_b, at low forward bias. Reverse bias stress causes recombination centers to form where the silicon-passivation oxide interface meets the emitter-base metallurgical junction. These centers add a recombination current with an ideality factor of two to I_b. Burnett and Hu [9] characterized the change in I_b with stress; Gummel plots of their fresh and degraded devices are shown in Fig. 3. Burnett and Hu measured ΔI_b, the change in I_b at a given forward bias, as a function of stress time t and stress reverse current I_r, and found that $\Delta I_b \propto I_r^{m+n} t^n$, where m and $n = .5$. Kapoor *et al.* [15] observed an increase in the base resistance and a decrease in the peak f_t with stress, but did not explain these changes.

Various types of defects have been invoked by scientists working with modern transistors to explain the observed behavior of device parameters. Scientists have reported the creation of positively charged electron traps [2], of recombination centers that anneal very quickly at room temperature [16] and of recombination centers that anneal upon application of forward bias to the emitter base junction [1]. Though MOS capacitor experiments have produced a catalog of the defects created in silicon dioxide by hot-carrier injection, the defects created in bipolar transistors have never been compared with defects identified in other structures. The above scientists claimed that the above defects are created in BJTs without proving that they exist and without explaining how they are created. In this thesis, we will identify the defects that are created during stress, show that their existence is physically reasonable, and describe how and under what conditions they are created.

Chapter 3

Prediction of Injected Carrier Type

The degradation of n-channel MOSFETs induced by channel hot-electron (CHE) stress is greatly influenced by the gate bias applied during stress. For example, M, the ratio of the substrate current I_{sub} to the source current I_s at a fixed gate voltage V_g and drain voltage V_d, decreases with stress if V_g during stress is near the threshold voltage V_t, and increases with stress if V_g and V_d during stress are nearly equal [17]. Scientists suspect that the gate voltage applied during stress controls the type of carrier injected into the gate oxide: hot holes injected into the gate oxide and trapped there during $V_g << V_d$ stress cause M to decrease, and hot-electron injection and trapping during $V_g \sim V_d$ stress cause M to increase [17]. Since the two different carrier types produce different defects, knowing which type of carrier is injected from a BJT's emitter-base junction into the overlying passivation oxide during reverse-bias stress is essential to understanding the hot-carrier degradation of BJTs.

Carrier heating and injection into SiO_2 during the CHE stress of MOSFETs and the reverse-bias stress of BJTs are similar. Electrons and holes gain high energy from the strong electric fields in either the drain-substrate junction of a MOSFET or the emitter-base junction of a BJT, then drift to either the gate oxide-silicon or passivation oxide-silicon interface and surmount the SiO_2 energy barrier. Various physical models of carrier heating and injection can be found in the MOSFET literature, yet we shall not use them to calculate and compare the rates of electron and hole injection in the experimental BJT (the models are described in detail in Chap. 4). Even if we could calculate both rates with perfect accuracy, we would still not know which hot-carrier type to ignore because the damage-creation probabilities of the two hot-carrier types are unknown. We will simply show that hole injection has no noticeable effect on MOSFETs except under very specific conditions. Those conditions don't exist in a reverse-biased, ungated emitter-base junction, thus we can ignore hole injection in the experimental BJT.

Hole injection from Si into SiO_2 is usually unobserved because the intrinsic energy barrier height for holes is 4.8 eV [18] and the barrier height for electrons is only 3.2 eV [19]. The injection rate of either carrier decreases exponentially with barrier height, thus during most hot-carrier stresses, injected electrons greatly outnumber injected

holes and are responsible for nearly all defect creation. The CHE stress of n-channel MOSFETs with $V_g << V_d$ is an exceptional stress because hot holes dominate defect creation [17]. This is possible because the drain-substrate junction electric field aids the motion of holes from the substrate to the Si-SiO_2 interface and on through the gate oxide, lowering the effective SiO_2 energy barrier faced by holes. A band diagram that shows the reduced barrier is shown in Fig. 4. Holes heated in the band minimum that drift to the interface without suffering collision face an energy barrier much less than the intrinsic barrier of 4.8 eV. The barrier is lowered by band-bending and the Schottky effect. Holes heated closer to the interface than the band minimum face an energy barrier lowered by the Schottky effect alone. The Schottky effect doesn't alter the electron energy barrier, though band-bending increases it.

Each stress of the experimental bipolar transistor was the application of a constant reverse current of 10 mA to the emitter-base diode. As the diode I-V curve in Fig. A3 shows, the stress current flows at $V_{be} \approx -5.1$ V and is a pre-avalanche current made up of carriers generated by band-to-band tunneling and impact ionization. Contour plots of the potential and the carrier generation rate at the emitter-base junction perimeter during stress are shown in Figs. 5 and 6 (The complete cross section of the experimental transistor and a description of the transistor simulator can be found in Appendix A). One can see that the electric field is almost parallel to the Si-SiO_2 interface near the generation region. The SiO_2 energy barrier faced by holes drifting from that region to the Si-SiO_2 interface is not reduced by either band-bending or the Schottky effect. In fact, holes that approach the Si-SiO_2 interface on the base side of the point of peak generation are impeded by the electric field. Since CHE stresses of n-channel MOSFETs show that hot hole injection is only significant when the SiO_2 hole energy barrier is reduced, the hole injection in the experimental BJT is negligible.

Chapter 4

Heating and Injection of Electrons

Free electrons in semiconductors rarely achieve the kinetic energy necessary for injection into SiO_2 because they suffer frequent inelastic collisions. Even in a very high electric field, the distance over which an electron must accelerate to gain the SiO_2 barrier energy is many times greater than the mean free path. Phenomena such as electron heating that involve ballistic motion over such distances are described as non-local. The semiconductor drift-diffusion equations, derived under the assumptions that local equilibrium exists between free electrons and the lattice and that carrier drift velocity is a function of local electric field, don't predict electron heating. Even an electron moving at the silicon saturation speed does not have enough kinetic energy to surmount the SiO2 energy barrier.

Boltzmann's equation contains terms that represent the ballistic motion of electrons and terms that represent the scattering and recombination events that keep electrons near equilibrium. Solutions to Boltzmann's equation properly describe non-local phenomena such as electron heating and can be obtained by the Monte Carlo method. Monte Carlo simulations are computationally complex and beyond the scope of this thesis. We choose to explain non-local heating in the emitter-base junction of BJTs with models that successfully predict similar heating in the drain-substrate junction of MOSFETs. Our aim is to predict the electron injection ratio, the ratio of the current of electrons injected into the passivation oxide above the emitter-base junction to the reverse current that flows simultaneously through the emitter-base diode.

Electron heating and injection into gate oxide take place during channel hot-electron (CHE) stress of MOSFETs. Scientists have predicted the ratio of the current of electrons injected into the gate oxide to the drain current in two ways: by calculating the electron temperature distribution and with the "lucky electron" method. An article by Hofmann *et al.* [18] offers a good example of the first approach. They assume that the flux of hot electrons injected into the gate oxide from a given point along the channel obeys the standard thermionic-emission formula with the oxide energy barrier, electron temperature and electron density as position-dependent parameters. The electron temperature at a given point in the MOSFET channel is equal to an integral along the channel of a function of the electric field, and therefore can indicate non-local heating. The oxide energy barrier is position-dependent because it

is modulated by the oxide electric field via the Schottky effect.

Tam *et al.* [19] used Shockley's "lucky electron" theory to predict the same CHE stress injection ratio mentioned above. A "lucky electron" is one that gains speed over long distances without scattering, then drifts to the gate oxide: it accelerates to high kinetic energy in the drain-substrate junction without suffering any collision, suffers an elastic collision that directs it towards the gate oxide, then avoids collision again in its flight to the oxide. The probability of an electron tracing such a trajectory is the product of the probabilities of the events that compose the trajectory. The fraction of the electrons at a given point along the channel that are injected into the gate oxide is equal to the sum of the probabilities of the "lucky electron" trajectories with all possible final kinetic energies.

The MOSFET electron heating theory that we choose to apply to the experimental BJT must be well suited to the electric field and electron density distributions that exist in its emitter-base depletion region during reverse-bias stress. As described in Chap. 3, each stress of the experimental transistor was the application of a constant reverse current of 10 mA to its emitter-base diode. The stress current is a pre-avalanche current made up of carriers generated by band-to-band tunneling and impact ionization. Contour plots of the potential, carrier generation rate and electron density during stress in the emitter-base depletion region at the emitter perimeter are shown in Figs. 5, 6 and 7. The ridge in the carrier generation rate contours lies along the emitter-base metallurgical junction, where the electric fields in the diode are highest. The sharp falloff in electron density near the point (.98,.99) indicates the boundary of the depletion region in the emitter. If the carrier generation rate were not so intense, the electron density would be very low ($< 10^5$ cm^{-3}) throughout the space-charge layer because $V_{be} \approx -5.1$ V and the base dopant concentration is very high. The shallow peak in electron density that exists within the depletion region is maintained by the strong rate of carrier generation.

Reverse-bias stress of BJTs is markedly different from CHE stress of n-channel MOSFETs in that electron density is not constant along the electric field lines in the region where electrons are heated, the emitter-base depletion region. As shown in Fig. 7, electron density increases steadily across the depletion region. The electron density contours are parallel to the equipotential lines shown in Fig. 5 since electrons are generated by the junction electric field. In an n-channel MOSFET, the concentration of electrons in the channel does not vary by orders of magnitude from source to drain; the electrons heated at the drain-substrate junction are supplied by the source. Joshi [8] found that the electron density gradient in the emitter-base depletion region of a BJT has an important effect. He writes that "transistors with tunneling emitter-base junctions are less susceptible to current gain degradation" than transistors with emitter-base junctions that avalanche at low reverse currents because tunneling electrons only appear in the conduction band at the emitter-side edge of the depletion region, where they can do no damage to the Si-SiO_2 interface.

Electron temperature calculations are most often performed when electron density variation is negligible, such as in n-channel MOSFETs. A few authors use the electron temperature equations of Takeda *et al.* to make those calculations. Equation 2 of Takeda *et al.* [21] is an energy conservation equation that includes electron density and

temperature as position-dependent parameters. Equation 3 is the simple relationship between electron temperature and electric field that results from equation 2 if the electron density is assumed constant. Takeda *et al.* and Hofmann *et. al.* [18] who model the gate current of n-channel MOSFETs, use equation 3, the equation in which electron density does not appear.

Burnett and Hu [22] use the electron temperature approach to calculate the electron injection ratio of BJTs. They account for the electron density gradient in the emitter-base depletion region in a simple yet crude way. They assume the electron concentration to be zero on the base side of the depletion region and a constant value on the emitter side. The step in electron density occurs 30 Å closer to the emitter than the metallurgical junction, where Burnett and Hu claim electrons "emerge from tunneling". They use equation 2 of Takeda *et al.* to calculate the electron temperature distribution based on the above assumption. The boundary condition they assume is that the electron temperature is zero at the step in electron density.

The "lucky electron" method handles the variation in electron density in a BJT's emitter-base junction elegantly and is well suited to calculate a BJT's electron injection ratio. The trajectory of every electron that makes up the stress reverse current in the emitter-base diode can be stated precisely. Each electron is generated at a specific point in the emitter-base depletion region by either band-to-band tunneling or impact ionization. Every electron is generated with nearly zero kinetic energy, then gains kinetic energy as it flows along an electric field line towards the emitter. Each electron either reaches the neutral part of the emitter or, if it is "lucky", enters the passivation oxide that overlies the emitter-base junction. Since the trajectories are so well defined, the probability of each trajectory can be easily formulated. Thus we use the "lucky electron" method of Tam *et al.* to calculate the electron injection ratio of the experimental BJT; the calculation is given below.

4.1 Calculation of the Electron Injection Ratio

Consider an electric field line ℓ in the emitter-base depletion region that crosses the peak of electron generation (see Fig. 8). Let x_o and x_f be the endpoints of ℓ in the base and emitter, respectively. Let x_i be a point on ℓ. Let $s(x_i)$ be the distance between x_i and x_o, $\psi(x_i)$ be the potential at x_i and $y(x_i)$ be the distance between x_i and the Si-SiO_2 interface. We wish to calculate P_i, the probability that an electron generated at x_i enters the passivation oxide. The electron follows the trajectory shown in Fig. 9. It drifts along ℓ from x_i to x_j and gains kinetic energy $K = \psi(x_j) - \psi(x_i)$, $K \geq \phi_b$, where ϕ_b is the SiO_2 barrier energy. Next the electron scatters elastically at x_j towards the Si-SiO_2 interface, maintaining enough momentum in the direction normal to the interface to surmount the SiO_2 energy barrier. The electron then drifts from x_j to the SiO_2. The probability that the electron follows that trajectory is

$$P_{ij} = \exp(\frac{-(s(x_j) - s(x_i))}{\lambda})\frac{ds}{\lambda_r}\frac{1}{2}(1 - \sqrt{\phi_b/K})\exp(-y(x_j)/\lambda),$$

where λ is the "scattering mean-free-path", λ_r is the "redirection scattering mean-free-path" and ds is the length of the differential segment of ℓ at x_j.

P_i is equal to the integral of P_{ij} over all x_j between x_i and x_f such that $K \geq \phi_b$. Let x_b be the point on ℓ such that $\psi(x_b) - \psi(x_i) = \phi_b$. Then

$$P_i = \int_{x_b}^{x_f} \exp(\frac{-(s(x_j) - s(x_i))}{\lambda}) \frac{ds}{\lambda_r} \frac{1}{2} (1 - \sqrt{\phi_b / (\psi(x_j) - \psi(x_i))}) \exp(-y(x_j)/\lambda).$$

P_i can be evaluated at any x_i by calculating the above integral along the electric field line that contains x_i. The electron injection ratio, R, is the weighted average of P_i over all x_i in the emitter-base depletion region. Let $G(x_i)$ be the electron generation rate at x_i. Then

$$R = \int P_i G(x_i) dV / \int G(x_i) dV.$$

Instead of taking the weighted average of P_i over the entire depletion region of the experimental BJT to predict its electron injection ratio, we simply calculate

$$R = \int P_i G(x_i) ds / \int G(x_i) ds$$

over all x_i on ℓ of Fig. 8, since ℓ crosses the region of peak electron generation and since the transistor is uniform along its width. According to Tam *et al.*, the values of λ and λ_r are 9.2 nm and 61.6 nm, respectively. As discussed in Chap. 3, the SiO_2 electron energy barrier in the experimental transistor is not reduced by the Schottky effect. The barrier height, ϕ_b, is thus equal to its intrinsic value of 3.2 eV. The value of R that results from the above constants and the data displayed in Fig. 10 is 10^{-6}.

The data of Fig. 10 are slightly inaccurate. The data are computed by FIELDAY, a semiconductor device equations solver. We did not tune FIELDAY's impact ionization model to correctly predict the multiplication of carriers in the emitter-base depletion region. The data of Fig. 10 were computed with the electron ionization rate set to zero. By comparing the I-V curve of the experimental emitter-base diode (see Fig. A3) to the I-V curve in Fig. 5 of Burnett and Hu [22], we estimate that M, the multiplication factor of our stress reverse current, is between 2 and 9. Neglect of impact ionization will effect calculation of R the most if M is high as possible and no electron generated by impact ionization enters the passivation oxide. In this worst case, R is reduced by a factor of 10 to 10^{-7}.

In Chap. 5, we derive from experimental data a value for R of 4×10^{-6}. The correspondence between the calculated and the measured values of R is good enough to show that the "lucky electron" model of Tam *et al.* can be applied to BJTs without changing the values of λ and λ_r.

Chapter 5

Defect Creation by Hot Electrons

A thorough understanding of hot-electron-induced defect creation in SiO_2 has been achieved through the study of hot-electron injection into the gate oxide of MOS capacitors. It has been difficult to determine which defects are produced by a given injection because defects in a gate oxide can't be directly observed. Almost all information about defects in a MOS capacitor's gate oxide comes from the capacitor's C-V and I-V curves. Shifts and distortions of these curves reveal changes in the oxide's charge state, but don't reveal the types of defects in the gate oxide. Defect types and defect-creation mechanisms have been identified by measuring C-V and I-V curve shifts under a variety of experimental conditions. The following experimental parameters have been varied: oxide electric field, oxide thickness, fluence of injected electrons, temperature, gate material and hydrogen content of the gate oxide. For example, Bright and Reisman [23] performed hot-electron injections over a broad range of oxide field strength and electron fluence. They concluded that at low fluence and low oxide field, the increase in negative charge in a MOS capacitor's gate oxide is due to trapping of hot electrons by native defects. At very high oxide field, positive charge accumulates in the gate oxide. This is believed to be due to holes generated in the gate oxide by impact ionization that get trapped on native hole traps [24].

In this chapter we describe the hot-electron stress experiments we performed on bipolar transistors, present the experimental data, and deduce from that data the defect types and populations created during stress. We briefly discuss the kinetics of defect creation. We did not perform experiments over a broad enough range of experimental conditions to produce a new model of hot-electron-induced defect creation in SiO_2. Instead, we show that the defects we observe can be explained by the existing model of defect creation in MOS capacitors. The defects created in the experimental BJTs are the same defects that are created in MOS capacitors with the same oxide field and thickness as the experimental transistor, subjected to the same fluence of injected electrons as calculated in Chap. 4.

5.1 The Defect-Creation Experiment

The experiments were performed at wafer level in a darkened Faraday cage, as described in appendix B. All the stress and measurement biases we applied to the emitter-base junction of the experimental bipolar transistor were applied with the transistor collector left floating. The collector is distant from the region of electron heating and is known to play no part in defect creation. Each stress was the application of a constant reverse current of 10 mA to the emitter-base diode for a set length of time at room temperature. The stress time ranged from 82 ms to 5000 s. As described in Appendix A, the stress current flows at the emitter perimeter and is made up of carriers generated by band-to-band tunneling and impact ionization. Many other researchers have used similar stress currents; our perimeter reverse stress current of .2 μA/μm is typical of previous experiments.

We measured the following three parameters to characterize fresh and degraded diodes: C_0, the junction capacitance at $V_{be} = 0$; I_{rec}, the current at $V_{be} = .3$ V; and I_{bb}, the current at $V_{be} =$ -1.9 V. I_{rec} is due to the recombination of carriers in the depletion region. It is measured at a forward bias high enough to avoid enhancement of the defect emission cross sections by the junction electric fields, but not high enough to cause the forward-bias defect annealing observed by other scientists (see Chap. 7). I_{bb} is due to band-to-band tunneling and is somewhat enhanced by the presence of midgap states. It is measured at a reverse bias low enough to avoid hot-electron damage during measurement.

Each experiment consisted of the following sequence of stresses and measurements. The three monitor parameters of a fresh device were measured before the application of reverse bias stress for a set time. The three parameters were then measured 50 times each over 5 minutes as the damaged device relaxed. Next, the device was subjected to a very short reverse bias stress lasting 82 ms. The three parameters were then measured 50 times over five minutes as the device relaxed for a second time.

Typical results of experiments with stresses shorter than 100 s are shown in Fig. 11. ΔC_0, the difference between the capacitance and its original value, is plotted, and not C_0, the total junction capacitance, because the difference is always less than one percent of the total. Fig. 12 shows typical results of experiments with stresses longer than 100 s. The most drastic difference between long and short stresses lies in the distinct relaxation behaviors of the recombination current. Both sets of plots show that the main effect of the very short current stress at $t = 300$ seconds is to bring the damaged diode back to its condition immediately following the first, much longer stress. The relaxation that follows stress is reversible, and not due to the permanent annealing of defects.

Figs. 11 and 12 display the raw data of two stress experiments. They give a graphic depiction of how an experimental diode relaxes following stress and how it responds to a short stimulus. The most significant data measured during each stress and relaxation sequence are the values of the three parameters before stress, immediately following stress and after the full 300 s of relaxation. With those data we can quantify the effects of stress and recovery. The data taken after the short current pulse are

not of interest because, as described above, the short pulse returns a damaged diode to its condition immediately following stress. The parameter transients that follow a pulse are the same as the transients that follow stress. The following table shows the symbols used for the nine significant data of each stress:

Table 5.1: Data symbols

	fresh	after stress	after stress and relaxation
Cap. at V_{be} = 0 V	$C_{0,o}$	$C_{0,s}$	$C_{0,r}$
I at V_{be} = -1.9 V	$I_{bb,o}$	$I_{bb,s}$	$I_{bb,r}$
I at V_{be} = .3 V	$I_{rec,o}$	$I_{rec,s}$	$I_{rec,r}$

Some quantities derived from those data are as follows:

$$\Delta C_{0,s} = C_{0,s} - C_{0,o}$$

$$\Delta C_{0,r} = C_{0,r} - C_{0,o}$$

$$\rho I_{bb,s} = I_{bb,s}/I_{bb,o}$$

$$\rho I_{bb,r} = I_{bb,r}/I_{bb,o}$$

$$\Delta I_{rec,s} = I_{rec,s} - I_{rec,o}$$

$$\Delta I_{rec,r} = I_{rec,r} - I_{rec,o}$$

The above six quantities gauge the damage induced by a single reverse-bias stress. To display the results of all our experiments, I plot the above quantities as functions of stress time (see Figs. 13, 14 and 15).

5.2 Identification of Defects

In the experimental transistor, defects are created in the passivation oxide that overlies the emitter-base depletion region. Recombination centers created on the Si-SiO_2 interface cause I_{rec} to increase with stress (see Fig. 15). As mentioned in the introduction, many authors have observed this increase in recombination current. The recombination centers are the same as the interface states created during the hot-electron injection of MOS capacitors.

Charged defects distort the shape of the experimental transistor's depletion region and alter the electric fields within it. The changes are evidenced by the shifts in C_0 and I_{bb} shown in Figs. 13 and 14. Since the base is more lightly doped than the emitter, charged defects cause greater distortion of the base side of the depletion region than the emitter side. Positively charged defects repel holes; they widen the depletion region and decrease the electric field at the metallurgical junction, decreasing C_0 and I_{bb}. Negatively-charged defects attract holes, raising C_0 and I_{bb}. Fig. 13 shows how $\Delta C_{0,s}$ and $\Delta C_{0,r}$ vary with stress time. Fig. 14 shows that $\rho I_{bb,s}$ and $\rho I_{bb,o}$ vary similarly. Positive charge seems to accumulate in the passivation oxide for at least the first 10 s of stress. Negative charge appears later and overwhelms the positive charge

after at most 1000 s. The amount of positive charge increases during relaxation after every stress.

The accumulation of positive charge alone could also explain the entire $\Delta C_{0,s}$ vs. stress-time curve. C_0 would rise if positive charge in the passivation oxide inverted a portion of the base in the space charge layer. The island of electrons in an inversion region would serve as a protrusion of the emitter into the base, effectively shrinking the distance between the emitter and base along the emitter perimeter. But this hypothesis does not explain the $\Delta C_{0,r}$ vs. stress-time curve, for if the amount of positive charge increases during relaxation, then if an inversion region existed between the emitter and base, relaxation would cause C_0 to increase. At long stress times, when $\Delta C_{0,s}$ is positive, C_0 decreases with relaxation. We are forced to postulate the existence of both positively-charged and negatively-charged defects in the passivation oxide.

The negative charges are easy to identify. They are electrons trapped in the passivation oxide that don't leak from the oxide during relaxation. The electrons that overwhelm the positive charge at high stress times are most likely trapped on stress-induced defects. DiMaria *et al.* [25] observed that great quantities ($> 10^{13}$ cm^{-2}) of traps are created in a MOSFET by a high fluence ($>$ .01 C/cm^2) of hot electrons. $\Delta C_{0,s}$ is small and positive after the shortest reverse-bias stress. This is probably due to a small quantity of native traps filled at low fluence. Bright and Reisman [23] observed that native electron traps in an IGFET are filled at much lower hot-electron fluence than stress-induced traps. There should be few native electron traps in the passivation oxide of the experimental transistor because the transistor is the product of a BiCMOS process and undergoes modern polycrystalline-gate processing. Such processing typically eliminates most water-related, native electron traps from SiO_2 [12].

The positive charges are more difficult to identify. The defects can't be trapped holes because, as we claimed in Chap. 3, significant hot-hole injection does not occur during reverse-bias stress. Trapped holes could not cause the decrease in C_0 with relaxation. The positive charges must be donor-type defects in the passivation oxide that discharge electrons into the silicon conduction band as a transistor relaxes, increasing the amount of positive charge in the passivation oxide. The defects refill during the pulse of reverse-bias stress. Fig. 16 shows how ΔC_0 relaxes after a 6 s stress. The slow decrease is typical of electrons tunneling from traps in SiO_2 into a silicon substrate. The logarithmic behavior indicates that electrons tunnel from traps distributed in the passivation oxide at various distances from the Si-SiO_2 interface [26]. The positively-charged defects must be similar to donor-type interfacial states, but removed from the Si-SiO_2 interface and only able to communicate with the silicon conduction band by tunneling.

The donor-type defects described above have been created in the gate oxide of MOS capacitors during hot-electron stress. They have been called "anomalous positive charge" (APC) [27] and "slow states" [28] and can be charged or discharged by biasing the gate to allow electrons to tunnel into or out of them. DiMaria and Stasiak [12] claim that APC, electron traps and interface states are all created via the same mechanism, which they call "trap creation". DiMaria *et al.* [24] write:

> Trap creation occurs when any electron with energy greater than 2 eV releases hydrogen from defect sites near the anode interface. This mobile species can then move to the cathode-oxide interface where it produces interface states and a distribution of oxide electron traps near this interface. This process is ...observed at fields as low as 1.5 MV/cm. It is measurable only after injection of $\geq .001$ C/cm^2 under positive gate voltages for structures that have undergone polycrystalline-silicon (poly-Si) gate processing.

The equipotentials in the passivation oxide and emitter-base depletion region of the experimental transistor during stress are shown in Fig. 5. The maximum electric field in the oxide is 1.5 MV/cm, the threshold field of DiMaria *et al.* It is likely that the positively-charged defects and the electron traps in the experimental transistor are the products of "trap creation". Possible trajectories of a hot electron in the passivation oxide and the hydrogen it releases are shown in Fig. 17.

Interface states contribute charge to the passivation oxide if they are empty, donor-type states or filled, acceptor-type states. In Chap. 7, we prove that the interface states are donor-type by showing that the annihilation of interface states is simultaneous with a decrease in positive charge. It is likely that the donor-type interface states and the "slow states" are related to each other. The interface states may be slow states that migrate to the Si-SiO_2 interface and achieve rapid communication with the silicon. As shown above, the slow states are distributed in the passivation oxide at various distances from the Si-SiO_2 interface. The interface states may be the slow states closest to the interface, the slow states that discharge trapped electrons the quickest. Indeed, Fischetti [28] noticed that fast, donor-type surface states in aluminum-gated MOS capacitors were generated by hot-electron injection at the same rate as slow states.

In Chap. 3 we claimed that hot-hole injection in the experimental transistor is negligible. The three defects identified above - donor-type interface states, trapped electrons and "slow states" - are all known to be created in MOS capacitors during hot-electron injection. Hot-hole injection is known to create interface states and trapped positive charge as well, but creation of the above three defects does not require hole injection. Our claim is thus confirmed by the experimental results.

5.3 Populations and Positions of Defects

We found the populations and positions of charged defects in the passivation oxide by simulating transistors damaged by hot-electron stress. We used computer simulation to calculate C_0 and I_{bb} for transistors with charged defects in different parts of the passivation oxide. We assume that the simulation of a transistor with defects represents an actual, stressed transistor when both share the same shifts in C_0 and I_{bb}.

The simulator we used was FIELDAY [5], IBM's semiconductor device equations solver. FIELDAY takes as input the finite-element model of the experimental transistor shown in Fig. A4. Fielday can assign fixed charge to any point in that model and

calculate C_0 and I_{bb} with those charges in place. The calculation is accurate because the two parameters depend only on the potential distribution in the emitter-base depletion region: C_0 depends on the width of the depletion region and I_{bb} depends on the electric fields in that region (FIELDAY uses the band-to-band tunneling model of Hurckx *et al.* [29]). For FIELDAY, finding the potential distribution is a simple electrostatics problem.

Reverse-bias stress of the experimental transistor produces islands of positive and negative charge in the transistor's passivation oxide. The positive charge is contributed by donor-type interface states and "slow states", and the negative charge is contributed by trapped electrons. We assume that the charge density on each island increases with stress, and that the islands don't grow in size during stress. During relaxation, the density of negative charge stays fixed because trapped electrons can't escape from the passivation oxide. The quantity of positive charge increases during relaxation due to discharge of electrons from donor-type slow states.

The number of possible positions and populations of two charged regions in the passivation oxide is enormous. We were able to limit the possibilities in a few ways.

- The amount of positive charge must not be enough to invert the base, for that would cause an increase in C_0 that is not observed, as described in the above section.

- DiMaria and Stasiak [12] claimed that slow states in gate oxides are located at the substrate-gate oxide interface, and that electron traps created by hot electrons are located in the oxide bulk, close to the interface. We followed DiMaria and Stasiak and placed the island of positive charge at the substrate-passivation oxide interface of the simulated transistor, and placed the negative charge in the oxide bulk.

- As shown in Fig. 17, hydrogen drifts towards the substrate-oxide interface from the top of the passivation oxide to produce both electron traps and slow states. For two reasons, the region of positive charge may extend further towards the left in Fig. 17 than the region of negative charge. First, hydrogen can be ionized and may drift towards the left in the oxide electric field. Since hydrogen drifts all the way to the Si-SiO_2 interface to create slow states, and creates electron traps in the oxide bulk, it would drift further to the left to produce slow states if it were ionized. Second, only filled electron traps contribute negative charge to the passivation oxide, thus the negative charge must be located close to the metallurgical junction, the source of hot electrons. The creation of positively-charged slow states by hydrogen doesn't require hot electrons, thus positive charge can be further from the metallurgical junction than trapped electrons.

We simulated transistors that had been subjected to four different amounts of stress and relaxation: 10 s of stress; 10 s of stress and 300 s of relaxation; 250 s of stress; and 250 s of stress and 300 s of relaxation. The experimental shifts in C_0 and I_{bb} of those transistors are tabulated below:

Table 5.2: Actual stress-induced shifts in transistor parameters

Condition	$\Delta C_{0,s}$ (fF)	$\Delta C_{0,r}$ (fF)	$\rho I_{bb,s}$	$\rho I_{bb,r}$
10 s stress	-175	-	-	-
10 s stress, relaxation	-	-453	-	-
250 s stress	+194	-	1.37	-
250 s stress, relaxation	-	-339	-	.96

The positions of the two charged regions in the transistor simulations that match the above data are shown in Fig. 18. The metallurgical junction meets the Si-SiO_2 interface at $x = .969$. The region of slow states extends along the interface from $x = .939$ to $x = .959$. The region of trapped electrons lies 19 Å above the interface and extends from $x = .955$ to $x = .964$. The densities of positive and negative charge in the two regions, and the calculated shifts in C_0 and I_{bb} that they produce, are tabulated below (The changes in I_{bb} are not fitted due to the effect discussed in Sec. 5.5).

Table 5.3: Simulated stress-induced shifts in transistor parameters

+, − charge dens. (cm^{-2})		$\Delta C_{0,s}$ (fF)	$\Delta C_{0,r}$ (fF)	$\rho I_{bb,s}$	$\rho I_{bb,r}$
$.7 \times 10^{12}$	0	-170	-	-	-
2×10^{12}	0	-	-430	-	-
$.5 \times 10^{12}$	-4×10^{12}	+219	-	1.40	-
2.5×10^{12}	-4×10^{12}	-	-342	-	.91

We have derived the populations of defects in the passivation oxide induced by two particular stresses and now wish to find the fraction of the stress reverse current that is injected into the oxide to create those defects. Fischetti [28] studied APC creation in MOS capacitors and found how the areal density of APC increases with hot-electron fluence. Fig. 2 of his article shows that a fluence of 40 mC/cm^2 creates an APC density of 2×10^{12} cm^{-2}. Our experimental data show that the same APC density is created by 10 s of reverse-bias stress. 100 mC of charge flow through the emitter-base diode of the experimental transistor during 10 s of reverse-bias stress at a constant current of 10 mA. Hot electrons are injected from the emitter-base depletion region into an area of oxide as long as the emitter perimeter, 5 cm, and approximately as wide as the island of positive charge described above, .02 μm; an area of 10^{-5} cm^2. The fluence of hot electrons injected during 10 s of stress is $R \times 10^7$ mC/cm^2, where R is the electron injection ratio. To match Fischetti's data, R must be 4×10^{-6}. This value of R, derived from experimental data, is only one order of magnitude greater than 10^{-7}, the value calculated in Chap. 4.

5.4 Defect-Creation Kinetics

Time did not allow me to match every stress and relaxation datapoint of Figs. 13 and 14 with a transistor simulation, thus we can't show how both the number of slow states and the number of trapped electrons in the experimental transistor's passivation oxide increase with stress. From a simple manipulation of the experimental data, we can

gain some understanding of the kinetics of the creation of slow states. The behavior of the experimental transistor following stress is controlled by only one type of defect: the slow states. We can monitor the population of slow states by comparing the transistor parameters before and after relaxation. The more the parameters change during relaxation, the more slow states are in the passivation oxide. Figs. 19 and 20 both indicate that the number of slow states grows for the first 200 s of stress, then saturates. Such saturation was observed by Fischetti [28]. He hypothesized that slow states are broken bonds at the Si-SiO_2 interface, and that their population saturates because only a few strained bonds at the interface can be broken during hot electron stress (the bonds are broken by the "mobile species" described in Sec. 5.2).

The number of fast interface states created during stress is generally assumed to be proportional to ΔI_{rec}. Fig. 15 shows that at low stress times, $\Delta I_{rec,s}$ and $\Delta I_{rec,r}$ are nearly proportional to a power of the stress time, which is typical of the creation of interface states in MOSFETs and previously measured I_{rec} degradation in BJTs. The behavior of $\Delta I_{rec,s}$ and $\Delta I_{rec,r}$ at high stress times is difficult to interpret. After long stresses, I_{rec} decreases substantially with relaxation, much as I_{bb} does. As discussed in Sec. 5.2, I_{bb} decreases because the electric fields in the emitter-base depletion region decrease with relaxation. Those electric fields must also control I_{rec} when large numbers of electrons are trapped in the passivation oxide. Electric fields above .5 MV/cm are known to increase recombination currents by enhancing the rate at which electrons are captured by interface states [30] (Field-enhanced recombination is fully discussed in Chap. 6). The $\Delta I_{rec,r}$ vs. stress time curve seems to indicate that interface state creation saturates after 100 s of stress; the increase in $\Delta I_{rec,s}$ with stress time after 100 s may be due to the electric field effect alone. One would expect the number of fast, donor-type interface states to saturate since, as described in Sec. 5.2, slow states and fast interface states are related defects.

Other authors have observed I_{rec} to saturate and to be enhanced by high electric fields. Hackbarth and Tang [16] found that after short reverse-bias stress, the forward current of their experimental emitter-base diode obeyed the standard formula for recombination current without field enhancement:

$$I_f = I_o \exp(qV_{be}/nkT),$$

with $n \leq 2$. They observed that I_o increased and then saturated in the early stages of stress. After long stress, the forward current obeyed the same formula, but with $n > 2$. Hackbarth and Tang claimed that the rise in n above 2 was due to electric field enhancement of the recombination current. Momose *et al.* [1] subjected BJTs to both high current and low current reverse-bias stress. They found I_{rec} of the heavily stressed transistors to decrease significantly during relaxation, and I_{rec} of the lightly stressed transistors to decrease only by a very small amount. Their data is very similar to mine; the behavior of I_{rec} that they observed can be explained by electric field effects.

We shall not discuss in detail the kinetics of electron trap creation and trap filling in our experimental transistor. In general, our experimental data agree with the data of others. DiMaria *et al.* [24] noted that electrons trapped in traps created by

hot-electrons are only observable after hot-electron injection of greater fluence than 1 mC/cm^2. As seen in Figs. 13 and 14, electron trapping in the experimental transistor becomes significant after 100 s of reverse-bias stress. In Sec. 5.3, we showed that the injected fluence after 10 s of stress is 40 mC/cm^2, thus the injected fluence after 100 s of stress is 400 mC/cm^2, above the threshold of DiMaria *et al.*

5.5 The Effect of Interface States on Tunneling Current

The description and analysis of our defect-creation experiments is nearly complete. We wish to remark once more upon the experimental data discussed in this chapter. In Sec. 5.2, we wrote that both Fig. 13 and Fig. 14 show the accumulation of positive charge in the passivation oxide in the early stages of stress, and the accumulation of negative charge in the late stages. The two figures are not identical; $I_{bb,s}$ is not less than $I_{bb,o}$ at low stress times, in spite of the positive charge in the passivation oxide. Some mechanism must compete with the decrease in electric field caused by the positive charge and make I_{bb} rise. We believe that this mechanism is trap-assisted tunneling, the tunneling of electrons from the conduction band to the valence band via the interface states discussed in Sec. 5.2. Such states are known to increase the current in an emitter-base diode current at all biases: the recombination current at forward bias, the generation current at low reverse bias and the trap-assisted tunneling current at high reverse bias. The population of the interface states doesn't change significantly during the relaxation that follows stress, thus the increase in positive charge with relaxation causes I_{bb} to drop.

Chapter 6

Field-Enhanced Recombination Current

When the electric fields in the depletion region of a p-n diode are high, tunneling enhances the rate at which electrons transit between the conduction band and the valence band. This rate enhancement causes the diode current in both forward and reverse bias to increase. In this chapter, we shall describe the increase of recombination current at low forward bias, and discuss the effect of high electric fields on the recombination current's activation energy, E_a. In Sec. 5.4, we claimed that high electric fields enhance the recombination current of the experimental transistor when electrons are trapped in its passivation oxide. In this chapter, we shall prove that claim by analyzing E_a of experimental transistors with and without trapped electrons.

When a diode is forward biased, an electron can transit from the conduction band to the valence band in different ways, each of which can be represented on an energy band diagram (see Fig. 21). Tunneling is represented by a horizontal line on a band diagram, because an electron loses no energy when it tunnels. The emission of phonons is represented by vertical lines, because phonon emission is energy loss. When there are no midgap energy states in the depletion region, the only allowed transition is purely vertical, the recombination of an electron and a hole accompanied by the creation of many phonons (see Fig. 21a). The total momentum of the emitted phonons is equal to the total momentum of the electron and the hole that recombine because momentum is conserved during this transition. Band-to-band recombination occurs rarely due to the momentum conservation requirement. No horizontal line connects the conduction band to the valence band, thus band-to-band tunneling is impossible.

Other transitions are possible when midgap states exist. Shockley-Read-Hall (SRH) recombination is one such transition. It is a purely vertical transition, the capture of an electron and a hole by a midgap state (see Fig. 21b). Momentum is not conserved during electron and hole capture because midgap states are localized [30], thus SRH recombination occurs more frequently than band-to-band recombination when the number of midgap states is moderately high. In a region of high electric field, electrons and holes that tunnel from their respective bands to midgap states increase the number of electrons and holes that recombine at those states, enhancing

the recombination current. Fig. 21c shows an example of field-enhanced recombination: an electron tunnels from the conduction band to a virtual energy state of the recombination center, then emits phonons and transits to the actual state, where it recombines with a captured hole. Field-enhanced SRH recombination is not hindered by a momentum conservation requirement because momentum is never conserved when an electron or hole transits to a localized state.

The current at low forward bias of a p-n diode with many midgap states, such as the emitter-base diode of a stressed experimental BJT, is dominated by both normal and field-enhanced SRH recombination. This dominance is obvious in the case of the experimental diode, because the diode's current at low forward bias is controlled by the population of midgap states on its passivation oxide interface, as discussed in Sec. 5.4. Since band-to-band recombination is not observed in the experimental diode, we shall exclude band-to-band recombination from the subsequent discussion.

Normal SRH recombination current is thermally activated. Most recombination occurs in the middle of the diode's depletion region, at the point where n and p are both equal to $n_i \exp(qV_f/2kT)$ (V_f is the forward voltage and all other variables have their usual definitions). Electrons in a diode's n-side and holes in the p-side must be thermally excited over potential barriers to reach that point. Normal SRH recombination current is proportional to the above n, thus the current's activation energy, E_a, is equal to $(E_g - qV_f)/2$ eV.

Field-enhanced recombination requires little thermal activation. As Fig. 21c shows, an electron that tunnels from the n-side of a diode to a midgap state reaches the middle of the depletion region without being thermally excited. Del Alamo and Swanson [31] found that field-enhanced recombination current is not thermally activated, but that it varies with temperature because E_g changes with temperature. Hurckx *et al.* [32] found that field-enhanced recombination current has a weaker dependence on temperature than normal SRH recombination current does.

Normal and field-enhanced recombination current occur simultaneously in a p-n diode, but E_a of the total recombination current reveals the degree to which the current is enhanced by high electric fields. If E_a is close to $(E_g - qV_f)/2$, field enhancement is negligible. If E_a is substantially less than $(E_g - qV_f)/2$, field enhancement is important.

In Sec. 5.4, we claimed that the recombination current of the experimental transistor is field-enhanced when trapped electrons in the passivation oxide raise the electric fields in the emitter-base depletion region. The rise in electric field also causes the rise in $\Delta C_{0,s}$ and the rise in $\rho I_{bb,s}$ shown in Figs. 13 and 14. To verify our claim, we measured E_a of a transistor that had been subjected to 6 s of reverse-bias stress and of another transistor that had been subjected to 1250 s of stress. Only the transistor that had been stressed for 1250 s contained many trapped electrons, because electron trapping is only significant after 100 s of stress (see Sec. 4.4). The experiment and the results are discussed below.

6.1 The Activation-Energy Experiment

All stresses and measurements were performed at wafer level in a darkened Faraday cage. Wafers were mounted on a Temptronix 3000 heat chuck, which can apply temperature with a precision of .1° C. Each stress was the application of a constant reverse current of 10 mA to the emitter-base diode of an experimental transistor for either 6 s or 1250 s at 25.0° C (see Sec. 5.1 for further description of reverse-bias stress). Following stress, the diode relaxed at 25.0° C. Relaxation was interrupted at various times to measure E_a.

To measure E_a, I_{rec} is measured at 25.0° C, then the temperature is quickly brought up to 45.0° C and I_{rec} is measured again. The temperature is then brought back down to 25.0° C. The total time needed to pulse the temperature is approximately one minute. The temperature is not raised high enough to cause substantial defect annealing. As long as E_a is not measured in the first five minutes of relaxation, the discharge of electrons that takes place during relaxation (see Figs. 11 and 12 and Sec. 5.2) does not significantly effect the value of E_a measured.

To calculate E_a, we assume that $I_{rec} \propto \exp(-E_a/kT)$. We extract E_a from the following formula:

$$E_a = k(\frac{T_{hi}T_{lo}}{T_{hi} - T_{lo}}) \ln(\frac{I_{rec,hi}}{I_{rec,lo}}),$$

where $I_{rec,hi}$ is the recombination current measured at 45.0° C, $I_{rec,lo}$ is the recombination current measured at 25.0° C, $T_{hi} = 318$ K and $T_{lo} = 298$ K. E_a is meant to quantify the dependence of the recombination current on temperature, and is not equal to the activation energy of any particular transition. The experimental data are tabulated below:

Table 6.1: Activation energy data

Stress (s)	Relaxation (min.)	E_a (eV)
6	9	.401
6	16	.401
6	60	.403
1250	8	.347
1250	16	.358
1250	60	.372

I_{rec} is measured at a forward voltage of .3 V, at which the activation energy of normal SRH recombination current is $(E_g - qV_f)/2 = .41$ eV. E_a of the transistor stressed for 6 s is very close to that value; E_a of the transistor stressed for 1250 s is much less, at all stages of relaxation. This indicates that the recombination current is indeed only field-enhanced when electrons are trapped in the passivation oxide. Both stress times are long enough to allow the creation of many slow states in the passivation oxide. Slow states are donor-type defects that discharge electrons during relaxation, which decreases the electric fields in the emitter-base depletion region (see Sec. 5.2). The decrease in fields that follows the 6 s stress does not change E_a, but the same decrease in fields after the 1250 s stress does push E_a

from its field-enhanced value towards the normal value. This must mean that field enhancement only occurs above an electric-field threshold, and that electron trapping pushes the electric fields in the depletion region above that threshold. Only if the electric fields are above the threshold immediately after stress does electron discharge from slow states increase E_a during relaxation. Schenk [30] modeled field-enhanced recombination current and found that an electric field threshold for enhancement does indeed exist, and is approximately .5 MV/cm. The electric fields in the experimental transistor's depletion region when $V_f = .3$ are near Schenk's threshold.

Chapter 7

Annihilation of Interface States

In Sec. 5.4, we described the accumulation of interface states, slow states and trapped electrons in the passivation oxide of the experimental transistor during reverse-bias stress. We have not yet discussed changes in the populations of those defects after stress. Many scientists have observed I_{rec} to decrease spontaneously after an emitter-base diode is stressed [1, 16]. I_{rec} has also been observed to decrease greatly when the diode is thermally annealed or forward biased [1, 33]. Explanations of this decrease have involved changes in defect populations: the annealing of interface states and the detrapping of trapped electrons. In this chapter, we describe the behavior of the experimental transistor after stress. We distinguish reversible phenomena, such as electron discharge from slow states, from irreversible phenomena, such as defect annealing and defect annihilation. We compare the defect annihilation we have observed to a similar phenomenon that occurs in MOS capacitors.

After an experimental transistor is stressed for a long time, C_0, I_{bb} and I_{rec} decrease during relaxation. The decrease is due to the decrease of the electric field in the emitter-base depletion region caused by the discharge of electrons from slow states. After a short stress, C_0 and I_{bb} decrease for the above reason, but I_{rec} increases. As proven in Chap. 6, I_{rec} is only influenced by the electric field when electrons are trapped in the passivation oxide, after 100 s of stress. Since the relaxation of I_{rec} following short stress is not influenced by electric fields, it shows the influence of defect annealing.

Fig. 22 shows the behavior of I_{rec} after a short stress. The rapid increase of I_{rec} that immediately follows stress is obviously caused by the quick discharge of electrons from slow states. If the slow states were recharged, I_{rec} would return to its original value. Fig. 11c shows that the I_{rec} transient that follows 2 s of stress is repeated when the slow states discharge after being recharged by a short injection of hot-electrons. The explanation of the reversible increase of I_{rec} is as follows: as shown in Fig. 18, the defects in the passivation oxide are over the base of the experimental BJT, where $p > n$. The discharge of electrons from slow states increases the amount of positive charge in the passivation oxide, repelling holes from the oxide and attracting electrons, which drives p closer to n. According to the Shockley-Read-Hall formula for recombination, wherever the product of n and p is constant, such as in the depletion region of a forward-biased diode, recombination is most frequent where n and p are

equal. Since the electron discharge makes p and n more equal near the interface states in the passivation oxide, I_{rec} rises.

The decrease in I_{rec} over long relaxation times is not related to either electron discharge or field decrease. It can only be due to annealing or annihilation of interface states. Huang *et al.* found that the decrease of I_{rec} is not thermally activated at temperatures less than 100° C, indicating that the decrease is not due to thermal annealing. They found that the decrease is caused by the forward bias applied to an emitter-base diode to measure I_{rec}. We believe that the decrease of I_{rec} shown in Fig. 22 has the same cause.

7.1 The Effect of Forward Bias on a Stressed Transistor

We investigated the effect of forward bias on the experimental transistor by interrupting the relaxation that follows a 2 s stress with a pulse of forward bias. The data are shown in Fig. 23. The key characteristics of the data are listed below:

- The forward bias pulse causes the I_{rec} curve to shift downwards and the C_0 and I_{bb} curves to shift upwards.
- The downward slope of the I_{rec} curve is much greater before the forward bias pulse than after.
- Both ΔC_0 and I_{bb} decrease rapidly following the forward bias pulse. The sharp drops in C_0 and I_{bb} that follow the pulse are not as great as the rapid drops that follow the reverse-bias stress.
- The slopes of the ΔC_0 and I_{bb} curves away from the sharply decreasing transients are the same before and after the forward bias pulse.

Our explanation of the data is based on the following model. During the forward-bias pulse, electrons are injected from the emitter into the base and into the emitter-base depletion region. The electron density near the defects in the passivation oxide increases (see Fig. 18). The defects in the oxide after a stress of 2 s are slow states and interface states. As noted in Sec. 5.2, a slow state and an interface state are similar donor-type defects: interface states are on the surface of the passivation oxide and communicate with the silicon quickly; slow states are removed from that surface and communicate with the silicon slowly via tunneling. During the forward-bias pulse, electrons near the passivation oxide charge a fraction of the interface states. Charging annihilates some of those interface states. A few electrons tunnel into the passivation oxide and charge the slow states that are closest to the Si-SiO_2 interface, just as electrons charge those states during hot-electron injection. The slow states deep in the passivation oxide are not affected by the pulse, because electrons can't tunnel to them within the pulse duration.

The above characteristics can be explained as follows:

- The I_{rec} curves shifts downwards and the C_0 and I_{bb} curves shift upwards because of the annihilation of the interface states. The I_{rec} curve shifts downwards because the number of recombination centers decreases. The C_0 and I_{bb} curves shift upwards because the amount of positive charge in the passivation oxide decreases when the positively charged, donor-type interface states disappear.

- The decrease of I_{rec} with relaxation time prior to the forward-bias pulse is due to the annihilation of interface states by the pulses of low forward bias used to measure I_{rec}. The pulse of high forward bias annihilates so many interface states that, after the pulse, the I_{rec} measurements annihilate few more. Thus I_{rec} is nearly constant following the forward-bias pulse.

- ΔC_0 and I_{bb} decrease rapidly following the forward-bias pulse because the slow states filled by electrons during the pulse discharge when the pulse is finished. These rapid decreases are the same transients that follow reverse-bias stress and short hot-electron injections. The drops in C_0 and I_{bb} that follow the 15 s forward-bias pulse are not as great as the drops that follow the 2 s reverse-bias stress because forward bias does not fill slow states as efficiently as hot-electron injection does.

- The slow states deep in the passivation oxide discharge electrons extremely slowly following stress. These states cause the slow decrease in ΔC_0 and I_{bb} over long relaxation times. ΔC_0 and I_{bb} decrease slowly both before and after the forward-bias pulse because forward bias doesn't affect these states.

If interface states states are annihilated when they are charged by electrons, they should be annihilated by a short pulse of hot-electron injection as well as by forward bias. To verify this, we performed a experiment very similar to the one described above, except that a 15 ms pulse of reverse-bias stress was substituted for the 15 s pulse of forward bias. A pulse of stress is more effective at filling slow states than a forward-bias pulse, as is evident from the ΔC_0 and I_{bb} transients that follow both pulses. This is because electrons are injected directly into the passivation oxide during reverse-bias stress. A pulse of stress is less effective at annihilating positively charged defects than a forward-bias pulse is; I_{rec}, ΔC_0 and I_{bb} are shifted more by a forward-bias pulse than by a pulse of stress. These results prove that interface states, not slow states, are annihilated by charging.

Though charging was not previously observed to destroy interface states, Kerber [34] and Trombetta *et al.* [27] did observe that slow states are annihilated when charged. Since the slow states and the interface states in the experimental transistor are related defects, one would expect charging to destroy both defects. We observe only interface states to be annihilated by forward bias and hot-electron injection; the reason for this is unknown.

We desired to prove that annihilation of interface states is caused by charging and not by any other process. To investigate the annihilation mechanism, we interrupted the relaxation that follows a 2 s stress with five different forward-bias pulses. The results of this experiment are shown in Fig. 24. The change in I_{rec} induced by each forward-bias pulse is tabulated below:

Table 7.1: Results of the interface-state annihilation experiment

$I_{forward}$	$V_{forward}$ (V)	ΔI_{rec} (nA)	accumulated ΔI_{rec} (nA)
1 μA	.472	-.7	-.7
10 μA	.559	-.9	-1.6
100 μA	.634	-1.16	-2.76
1 mA	.726	-1.49	-4.25
10 mA	.894	-1.52	-5.77

Lang and Kimmerling [35] discovered that defects in GaAs are annealed when electrons and holes recombine at those defects. They found that the annealing rate is proportional to recombination current. The same mechanism can not be responsible for the annihilation of interface states that we have observed, because ΔI_{rec} is not proportional to $I_{forward}$.

The most prominent feature of the above data is the linear dependence of accumulated ΔI_{rec} on $V_{forward}$ (see Fig. 25). It is difficult to prove that this dependence is caused by the annihilation of interface states by charging. An interface state is filled at a given forward bias if its energy is lower than the hole quasi fermi energy, or if its energy is between the electron and hole quasi fermi energies and it is in a position where $n > p$. The interface states are distributed on the passivation oxide surface and may be distributed in energy as well. The fraction of the interface states that are filled by electrons at a given forward bias depends on these distributions, which are unknown. It is likely that distributions exist for which the fraction of filled interface states increases linearly with forward bias. These distributions would produce the curve shown in Fig. 25 if interface states are indeed annihilated when filled.

Another feature of the data is that the total change in I_{rec} induced by all five pulses, -5.77 nA, is very close to -5.41 nA, the change in I_{rec} induced by the forward-bias pulse of the first experiment. The pulse of the first experiment was $V_{forward} =$ .8 V for 15 s. Apparently, the decrease in I_{rec} caused by forward bias saturates; I_{rec} can not be returned to its pre-stress value by forward-bias annealing. Huang *et al.* [33] observed the same saturation.

Chapter 8

Conclusion

This thesis gives the most precise microscopic description to date of the hot-electron degradation of bipolar transistors. A description of such detail should help reliability engineers to understand stress-induced degradation of their bipolar transistors, and to establish the reliability of any BJT or BiCMOS technology. The heating of electrons in the emitter-base depletion region of the experimental transistor is described with the "lucky electron" model. The three defect types created by hot electrons are identified to be donor-type interface states, donor-type "slow states" and trapped electrons. The same quantities of the same defects are created in MOS capacitors under similar stress conditions.

The influence of each defect on transistor parameters such as base current is explained in this thesis. We have duplicated the defect-induced changes in parameters of stressed transistors with computer simulations of transistors with defects. In this way, we have found the populations and positions of the defects in damaged devices.

In this thesis we explain why transistor parameters change as a stressed transistor relaxes following stress. The discharge of electrons from slow states during relaxation, which lowers the electric fields in the emitter-base depletion region, causes the emitter-base junction capacitance and the band-to band tunneling current to decrease. The same process causes the base current to decrease when that current is produced by electric-field-enhanced recombination. The base current also decreases when interface states are annihilated during forward bias of the emitter-base junction.

This thesis draws together the study of hot-electron effects in MOS devices and the study of the same effects in bipolar transistors. The latter study gains the most from the union, for detailed models of electron heating and defect creation have resulted from the former. The models of electron heating and defect creation in BJTs included in this thesis are based on existing models of the same events under similar conditions in MOS devices to make the BJT models as precise and valid as possible.

Designs of bipolar transistors with reduced vulnerability to hot electrons have been proposed by various authors. In general, to increase the lifetime of a bipolar transistor, the electric fields in the emitter-base depletion region are lowered, thereby lowering the number of hot electrons in that region. The design of hot-electron-resistant BJTs is not discussed in this thesis because the models of electron heating and defect creation described in it can be applied to bipolar transistors of all designs.

There are two research topics that could be studied to further improve the understanding of hot-electron degradation of bipolar transistors. The first is defect creation by hot-electrons during low reverse-bias stress. It was claimed in Sec. 5.2 that no slow states or electron traps would be created in the passivation oxide of the experimental transistor if the electric fields in that oxide were weaker, that is, if the transistor were stressed at a low reverse bias. It remains to be seen whether or not slow states or electron traps would be created at low reverse bias. If those defects are created at lower oxide fields, then the existing theory of defect creation in silicon dioxide would be called into question. The second area of research is the annihilation of interface states. It was claimed in Sec. 7.1 that interface states are annihilated when they are charged. The physics and chemistry of annihilation remain to be examined. MOS devices would be the best experimental vehicle for study of the annihilation of interface states.

Appendix A

The True and the Simulated Experimental Bipolar Transistor

Since hot-carrier effects in a BJT don't involve the collector, only the emitter-base junction of the experimental transistor are described below, and only its emitter and base are simulated.

The experimental transistor is the product of a BiCMOS process (see Fig. A1). The process includes the formation of subcollectors as well as n-wells, thick oxide isolation and gate oxide identical to those produced in the normal CMOS process prior to the introduction of other bipolar elements. A thin polysilicon film is deposited immediately on the freshly grown gate oxide to protect it from subsequent processing. A mask then defines the base windows, inside which the first poly layer is removed and a passivation oxide called the base oxide is grown. The base oxide is grown in a manner similar to high-quality gate oxides. The base is then implanted through the base oxide at low energy to keep the width of the base small. Windows for emitter diffusion are opened in the base oxide just prior to the second polysilicon deposition. This polysilicon forms the emitter and completes the polysilicon gate stack. The second polysilicon is implanted with arsenic in order to achieve a shallow (approximately 100 nm) emitter. A blanket silicon nitride film is deposited over the polysilicon at this point in the process and the entire stack is etched during the gate and emitter definition. A low temperature oxide spacer is deposited and etched, following which the NFET and PFET source/drains are implanted. The PFET source/drain forms the base contact of the npn bipolar transistor and is confined to the silicon nitride overlayer of the emitter polysilicon stack. This nitride layer is removed with hot phosphoric acid prior to forming titanium silicide. The BiCMOS process concludes with the laying of local interconnections and metallization, just as in the normal CMOS process.

The resultant experimental transistor has a base oxide thickness of 340 Å. The transistor's SIMS profile, measured through the emitter window in the base oxide, is shown in Fig. A2. The silicon wafer surface is indicated by the high concentration of arsenic at the polysilicon-silicon interface. Notice that the peak boron concentration is quite high (approximately 2×10^{18} cm^{-3}) and is only 300 Å below the wafer surface. The high base doping makes the electric fields in the emitter-base space charge region

high as well. The depth of the boron peak is less than the emitter depth, thus the highest electric fields in the emitter-base junction occur just under the base oxide, along the emitter perimeter, where the peak of boron meets the arsenic emitter diffusion. That high-field region is the site of band-to-band tunneling, impact ionization, junction breakdown and carrier heating.

The experimental transistor is designed in the shape of a comb with long, narrow teeth and an enormous 5 cm perimeter to maximize the population of hot carriers. An I-V plot of the emitter-base diode is shown in Fig. A3. The bow in the leakage current curve is the tell-tale sign of band-to-band tunneling. The tunneling current is higher than the noise level even at very low reverse biases because the tunneling occurs in the high electric field region all along the emitter perimeter (We verified that the leakage current flows at the emitter perimeter by comparing the leakage current of emitter-base diodes of various sizes). The low breakdown voltage is indicative of the experimental transistor's high base doping.

It is desirable to make a finite-element model of the experimental emitter-base junction to simulate the effect of trapped charge in the base oxide on diode I-V and C-V characteristics. We use the IBM process simulator FEDSS [4] to construct our two-dimensional model diode and the IBM semiconductor device equations solver FIELDAY [5] to calculate its capacitance and current at various biases. The substrate was assumed to be a lightly and uniformly doped with boron at the start of the FEDSS process simulation. The simulation proceeds with deposition of the 340 Å base oxide, base implantation, etching of the emitter window in the base oxide and deposition of the emitter polysilicon doped *in situ* with arsenic. The final step in the simulation is a heat cycle that drives in the implanted base and the emitter.

The process simulation produces a model of the emitter-base diode (see Fig. A4) that is one half of the true, axially symmetric diode: the emitter window width is half the true .8 μm emitter width and the width of the polysilicon that overhangs the emitter perimeter is 1 μm, a dimension taken from the polysilicon-etch mask. Ohmic contacts to the model diode are placed on the polysilicon emitter and on the left edge of the base, where the p+ base contact would be. The main difference between the model device and the actual device is the vertical contour left by the etch of the emitter window in the base oxide. During the actual, isotropic etch, acid undercuts the etch resist and leaves a quarter-circle contour in the base oxide. The etch is simulated by the removal of a rectangle of SiO_2 from the base oxide, as seen in Fig. A4. Since we are interested in small changes in junction capacitance induced by charge in the base oxide, the shape of the oxide contour near the emitter perimeter is of some importance. Fortunately, the arsenic diffuses away from the corner of the oxide, diminishing the impact of the oxide contour. That the simulated emitter is *in situ* doped and not implanted is not significant because arsenic diffuses much more readily in polysilicon than silicon.

It is impossible to produce a discretized double of the experimental transistor simply by feeding the BiCMOS process recipe into the FEDSS program; a good match can be obtained only by optimizing the process simulation. Since the base doping profile has very strong influence over the emitter-base junction electric fields and thus the tunneling currents and the junction capacitance, we chose to match the

simulated emitter-base diode to the true diode by varying the base implant energy in the process simulation. We matched the C-V curves of a simulated and a real base oxide MOS capacitor, then simply added the emitter window etch to the process simulation to allow the arsenic to diffuse into the base. The comparison of the true diode I-V and C-V curves to the curves calculated by FIELDAY with the diode model (see Fig. A5) shows that the model diode is an accurate mimic (FIELDAY calculates the band-to-band tunneling current with the formulas of Hurckx *et al.*).

Appendix B

Experimental Setup

Fig. B1 shows the arrangement of equipment used to probe the experimental transistors at wafer level. A vacuum chuck keeps the wafer in place inside a darkened Faraday cage with a large felt portal. The cage is mounted on a vibration stabilizer and the wafer is connected to the outside world by Micromanipulator needles and by BNC cables that protrude the cage portal. Outside the cage stand the instruments that comprise the measurement circuit: an HP 59307 relay switch, a Kiethley 236 SMU and an HP 4284A precision LCR meter. The switch is connected directly to the BNC cables at the cage portal. The LCR meter is connected to the switch by 1 m of BNC cable and the SMU is connected to the switch by 1 m of triax shielded cable. Both meters are at ground when switches are performed. Each instrument is attached to an HP interface bus and controlled by an IBM PS/2 that runs command programs written in BASIC.

Figures

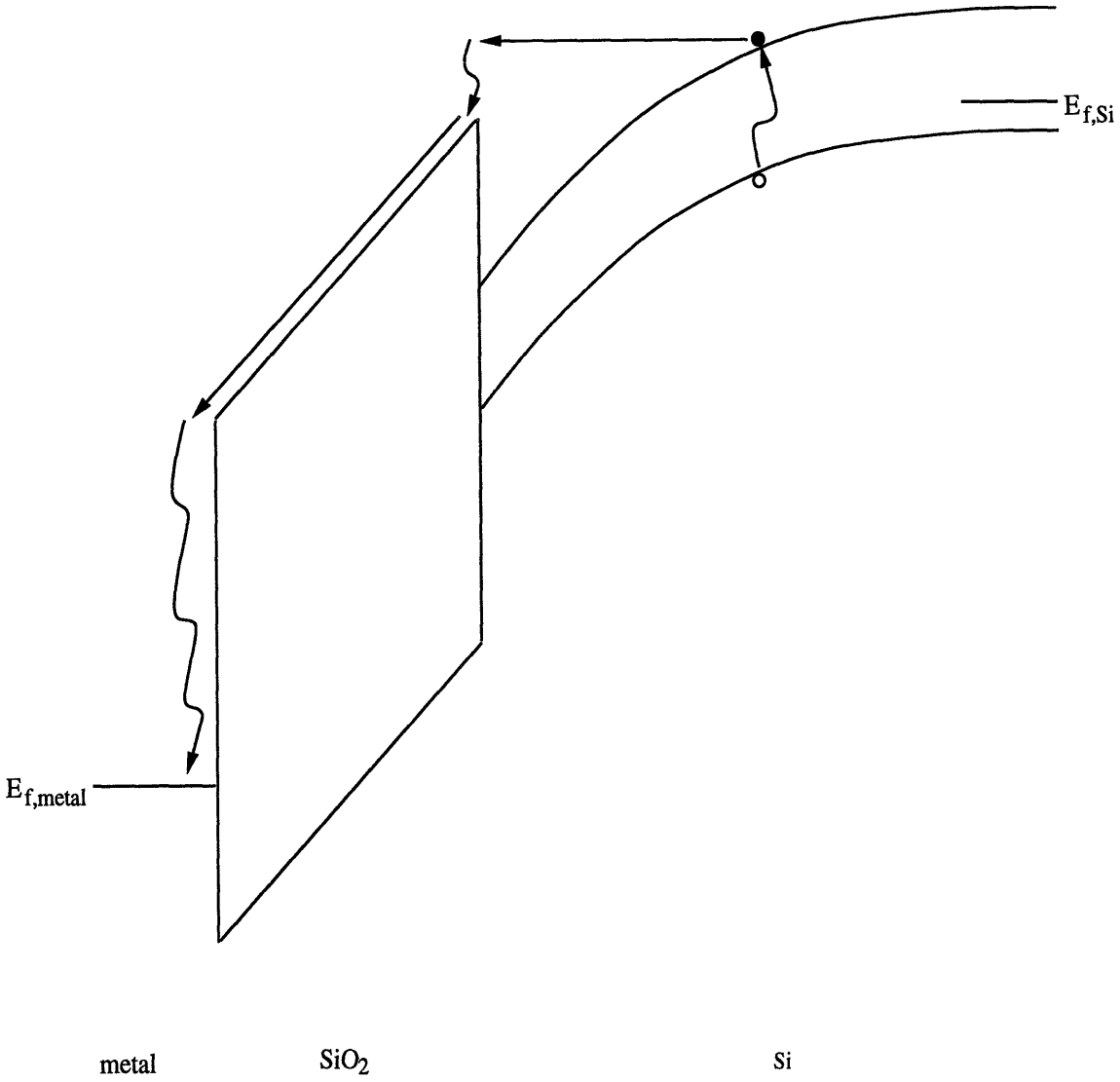


Fig. 1 Avalanche injection of electrons in a pMOS capacitor [1].

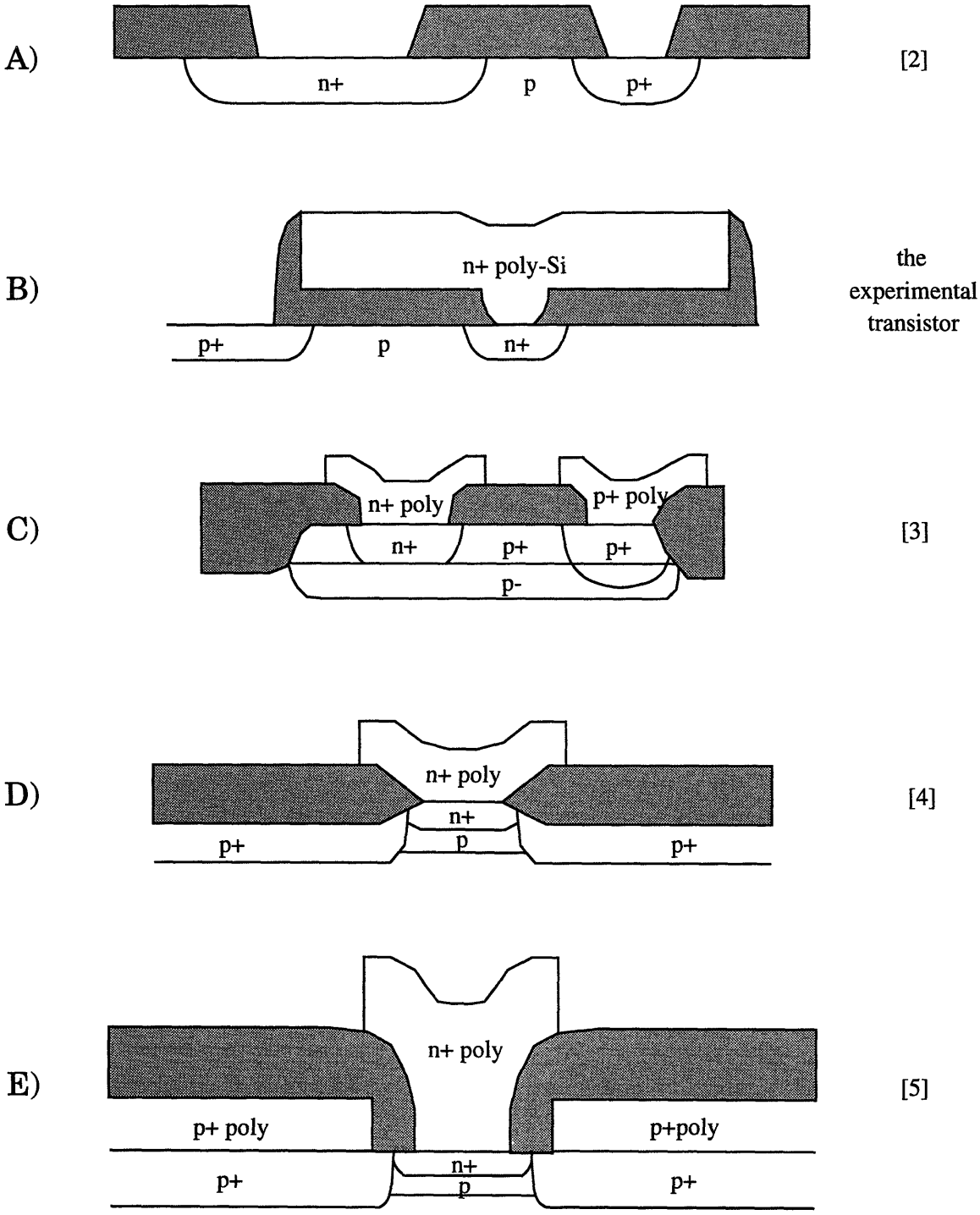


Fig. 2 Emitter-base diodes of past and present hot-electron experiments.

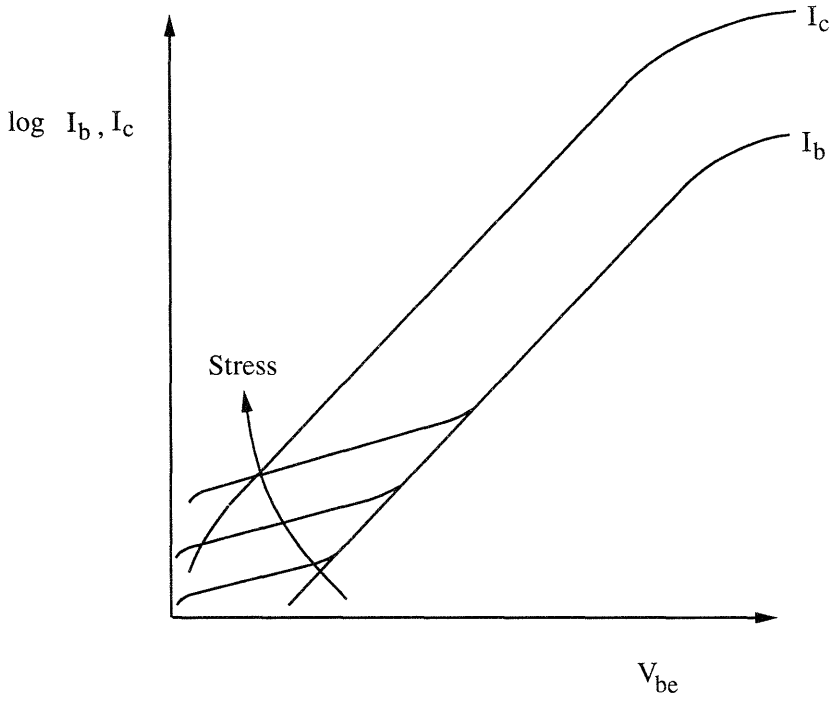


Fig. 3 Gummel plots of Burnett and Hu [9].

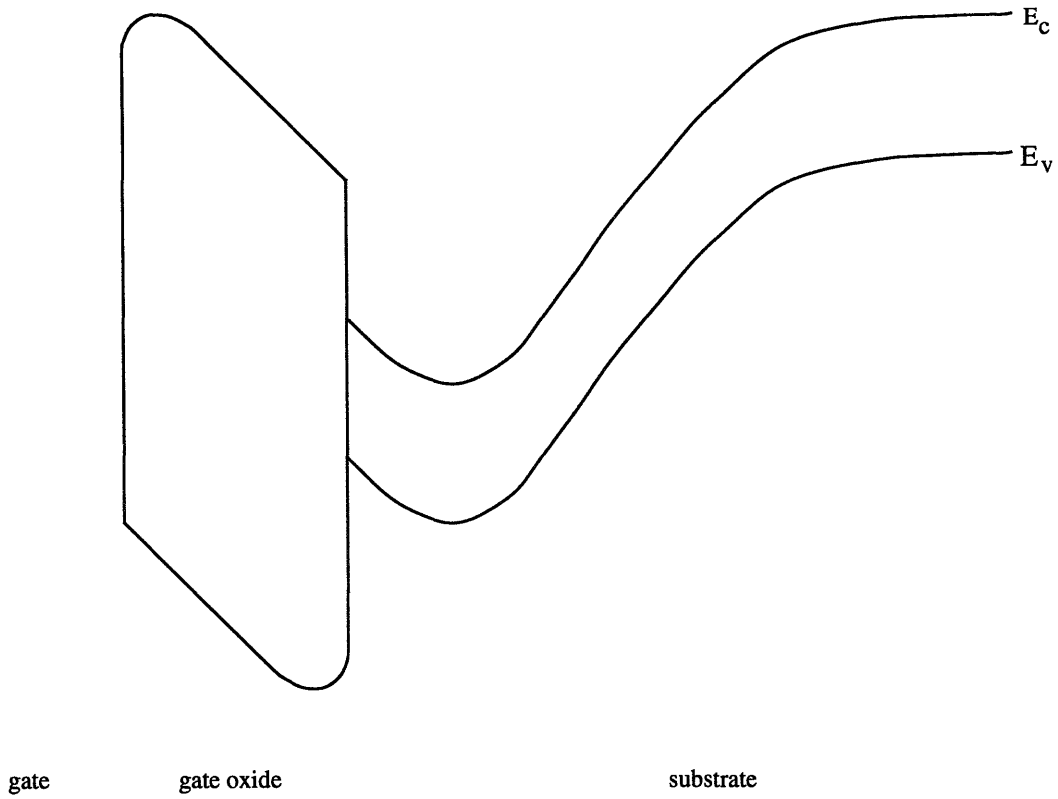


Fig. 4 Cross section of the energy bands near the drain of an n-channel MOSFET under $V_g << V_d$ stress. The cross section is made normal to the gate oxide [20].

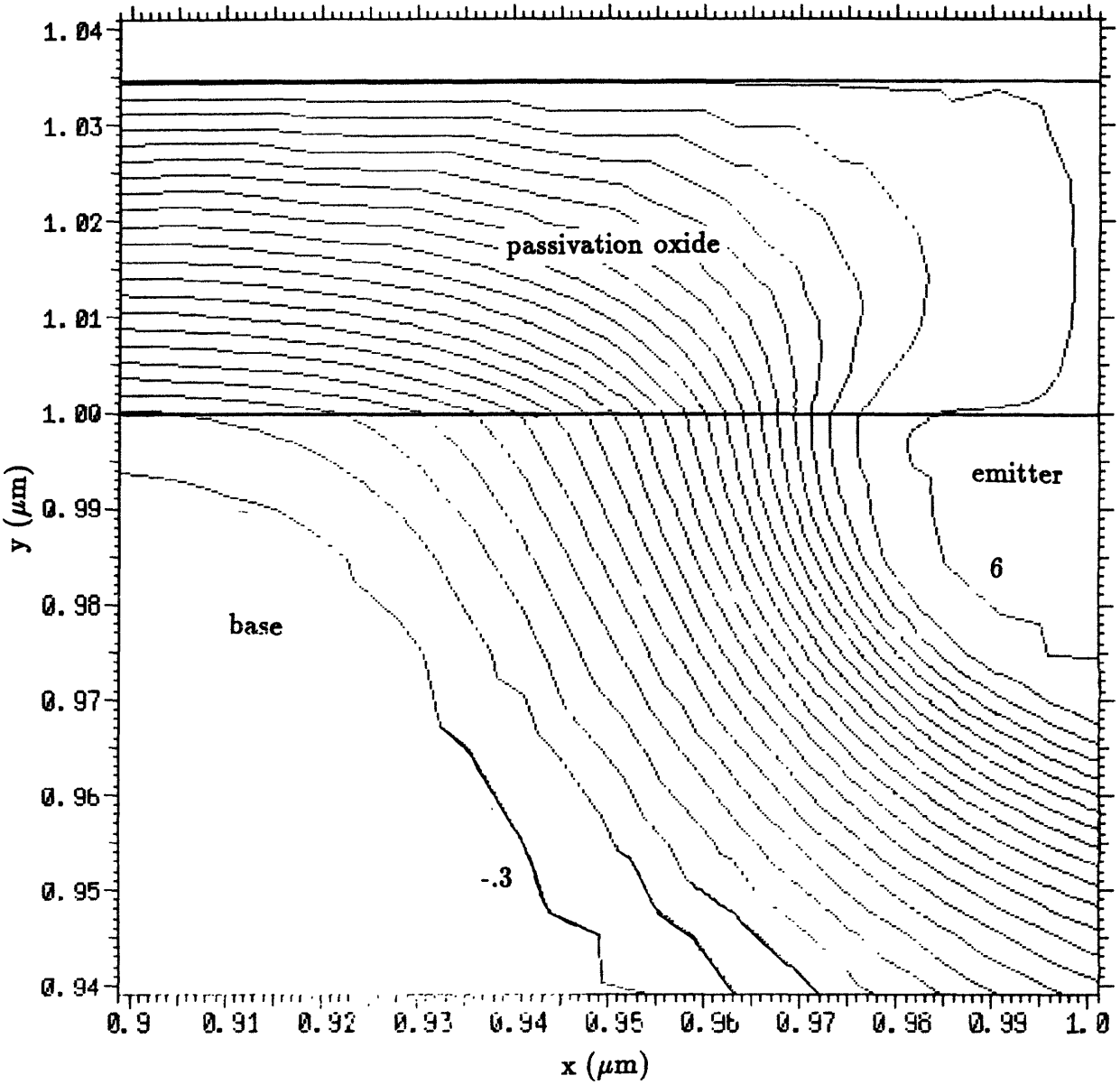


Fig. 5 Potential contours near the emitter perimeter during reverse-bias stress. The values of the potential on the contours run from -.3 V to 6 V in .3 V intervals.

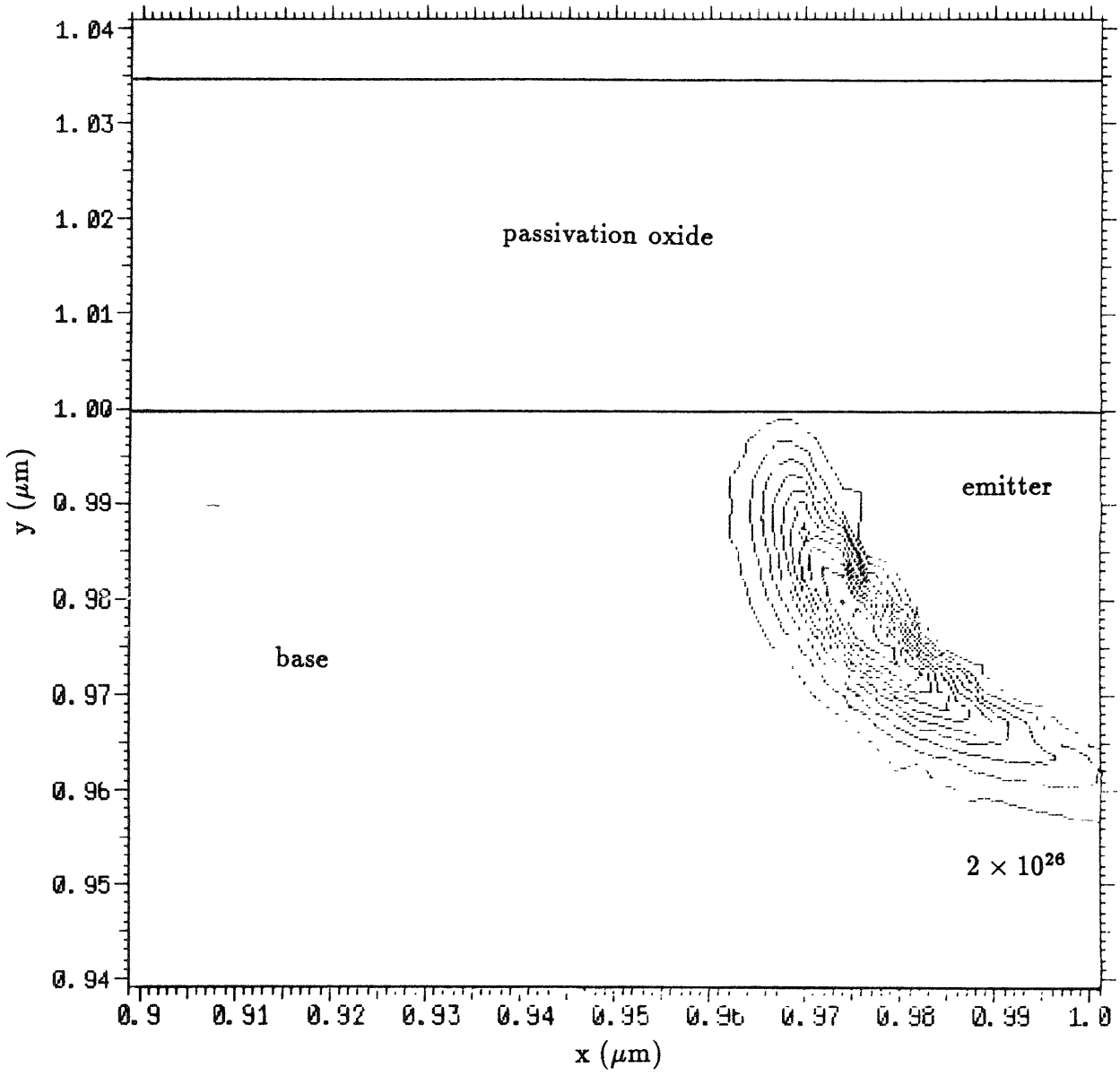


Fig. 6 Generation rate contours near the emitter perimeter during reverse-bias stress. The values of the generation rate on the contours run from 2×10^{26} $cm^{-3}s^{-1}$ to 2×10^{27} $cm^{-3}s^{-1}$ in 2×10^{26} $cm^{-3}s^{-1}$ intervals.

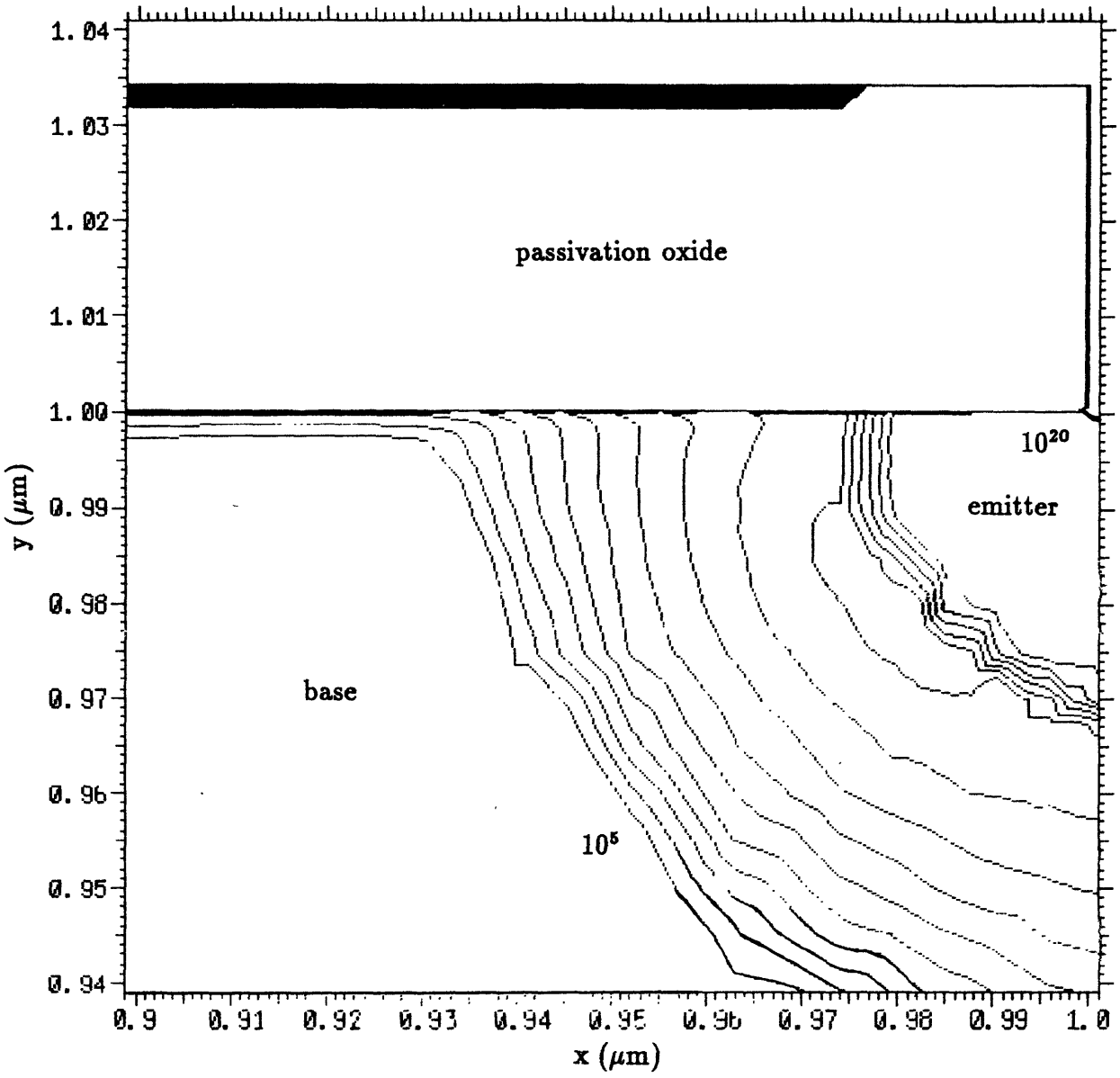


Fig. 7 Electron denstiy contours near the emitter perimeter during reverse-bias stress. The values of the electron density on the contours run exponentially from 10^5 cm^{-3} to 10^{20} cm^{-3} in intervals of one order of magnitude.

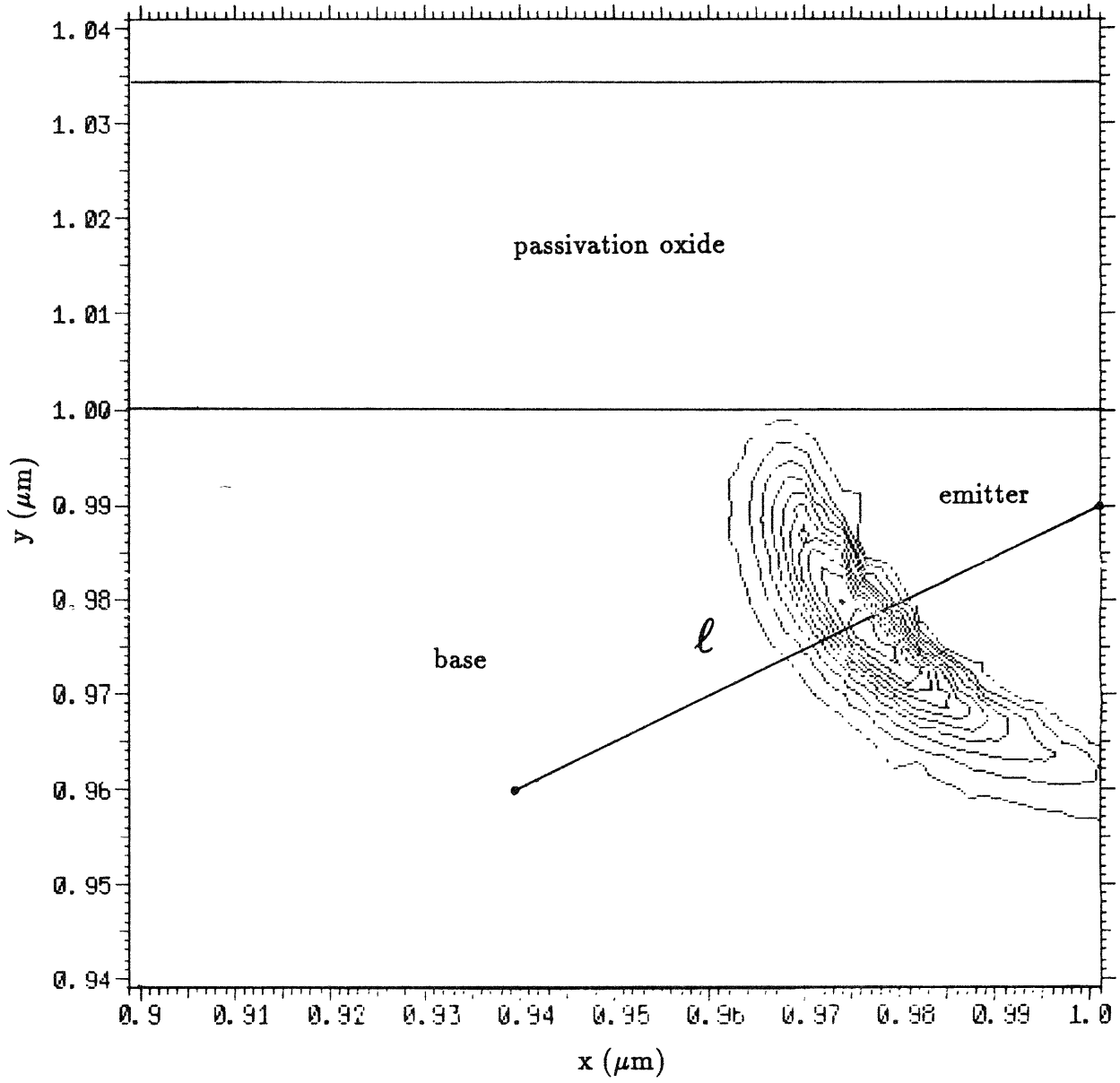


Fig. 8 The electric field line ℓ, which crosses the peak of electron generation (see Fig. 6).

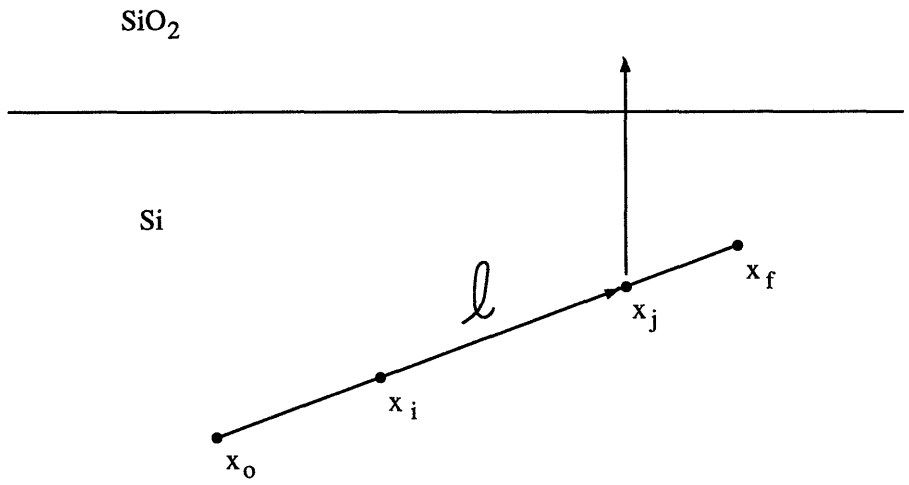


Fig. 9. Trajectory of an electron generated at x_i and injected into the silicon dioxide.

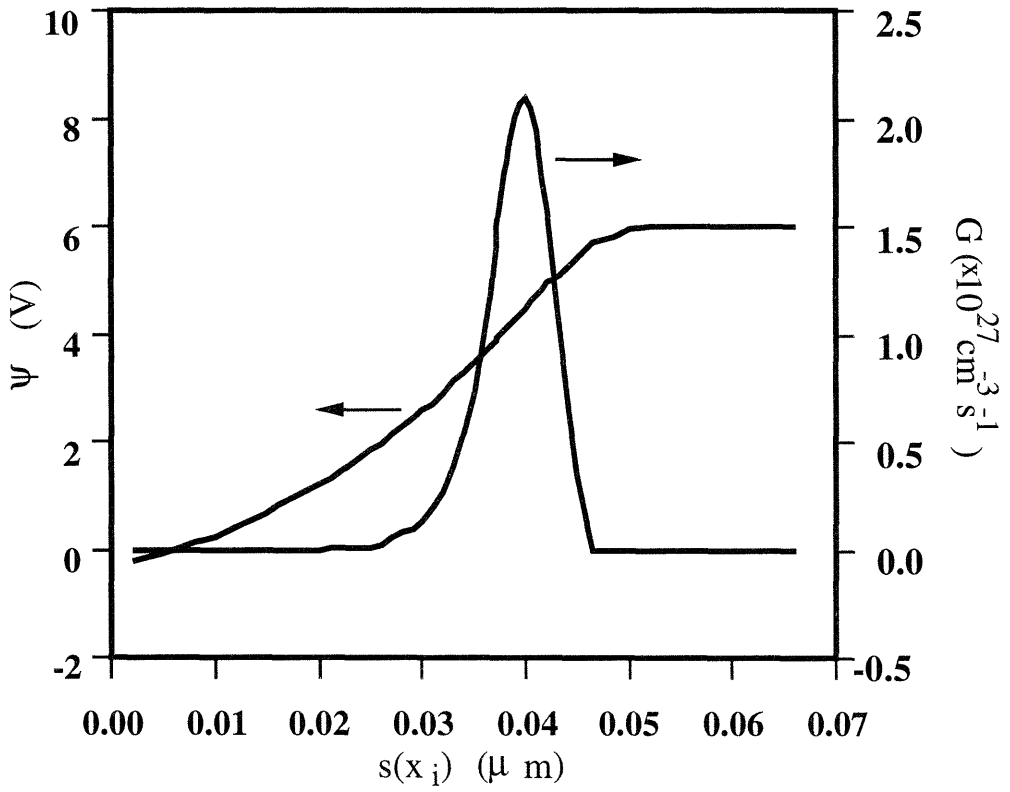


Fig. 10. Potential and generation rate at x_i on ℓ vs. $s(x_i)$.

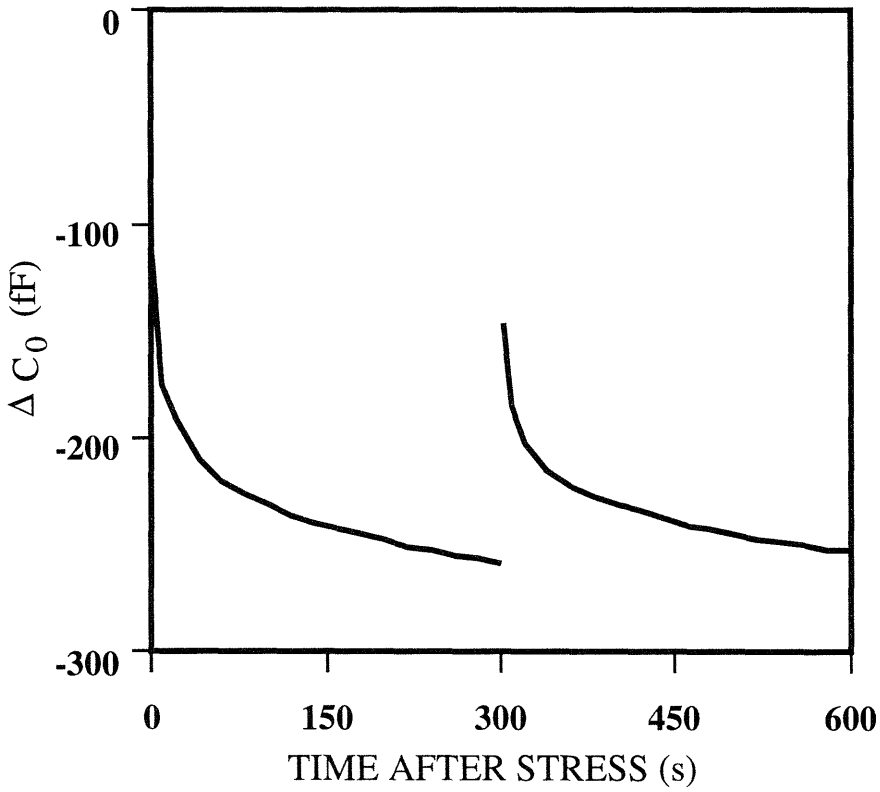


Fig. 11a. ΔC_0 after a 2 s stress.

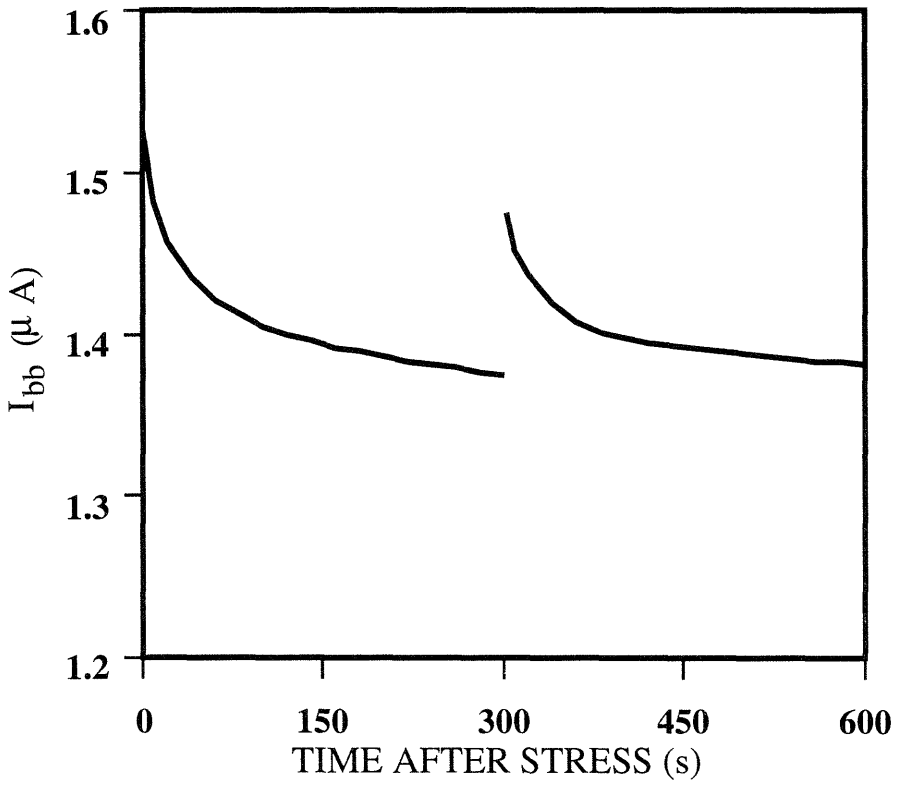


Fig. 11b. I_{bb} after a 2 s stress.

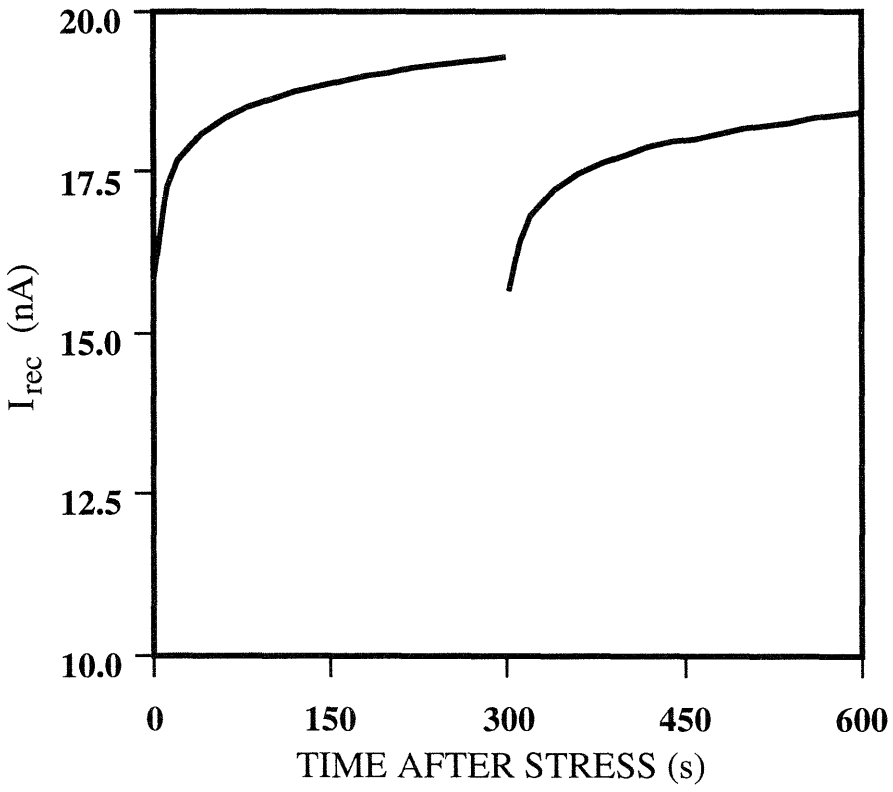


Fig. 11c. I_{rec} after a 2 s stress.

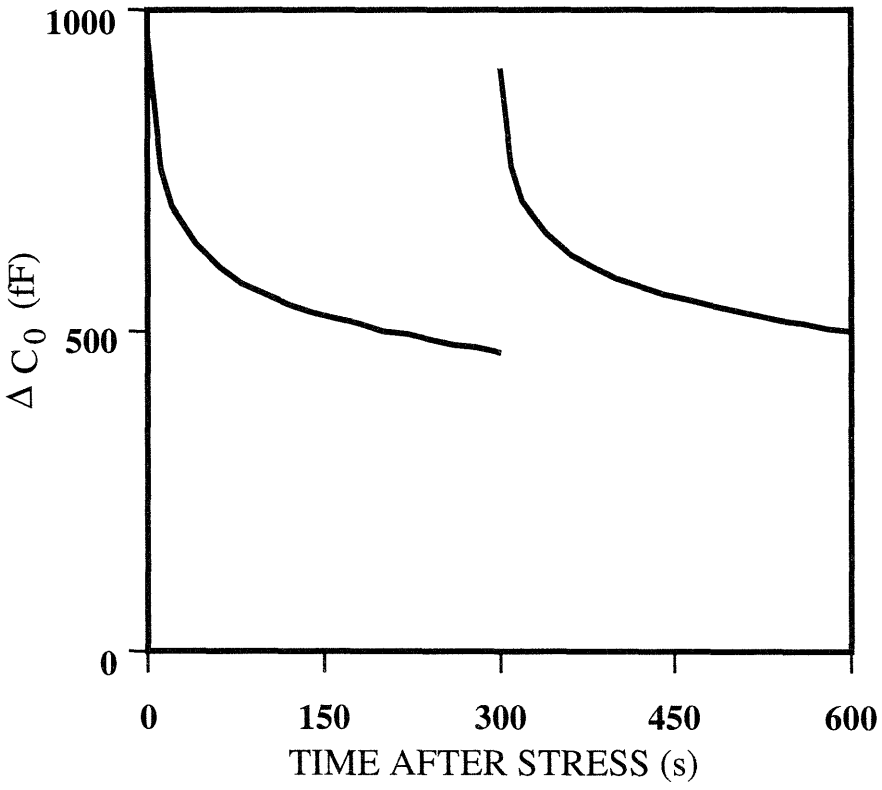


Fig. 12a. ΔC_0 after a 5000 s stress.

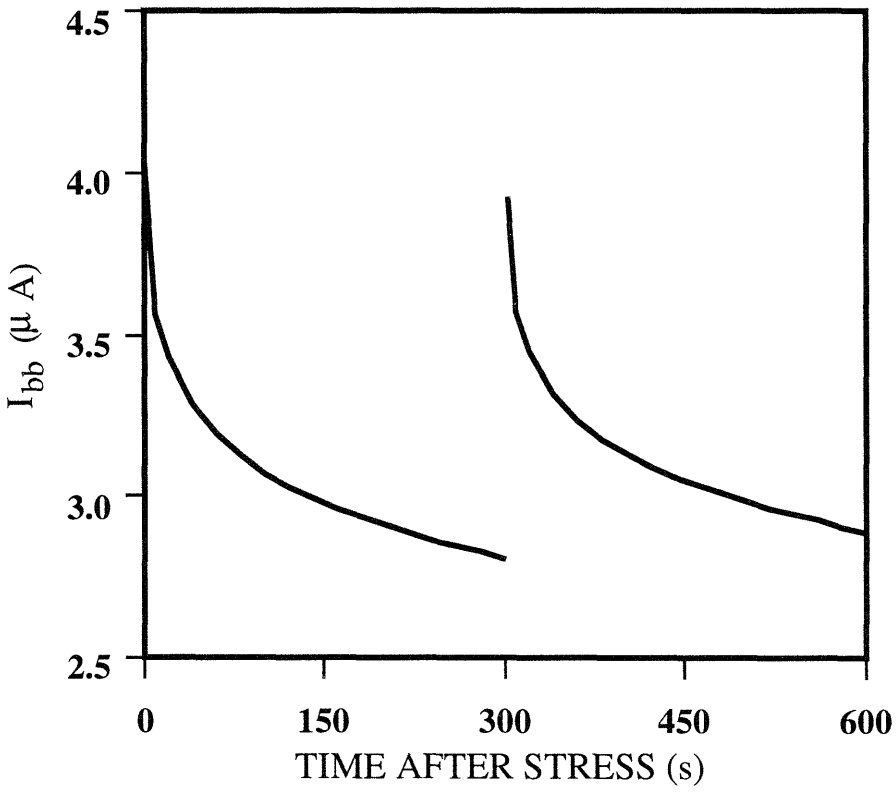


Fig. 12b. I_{bb} after a 5000 s stress.

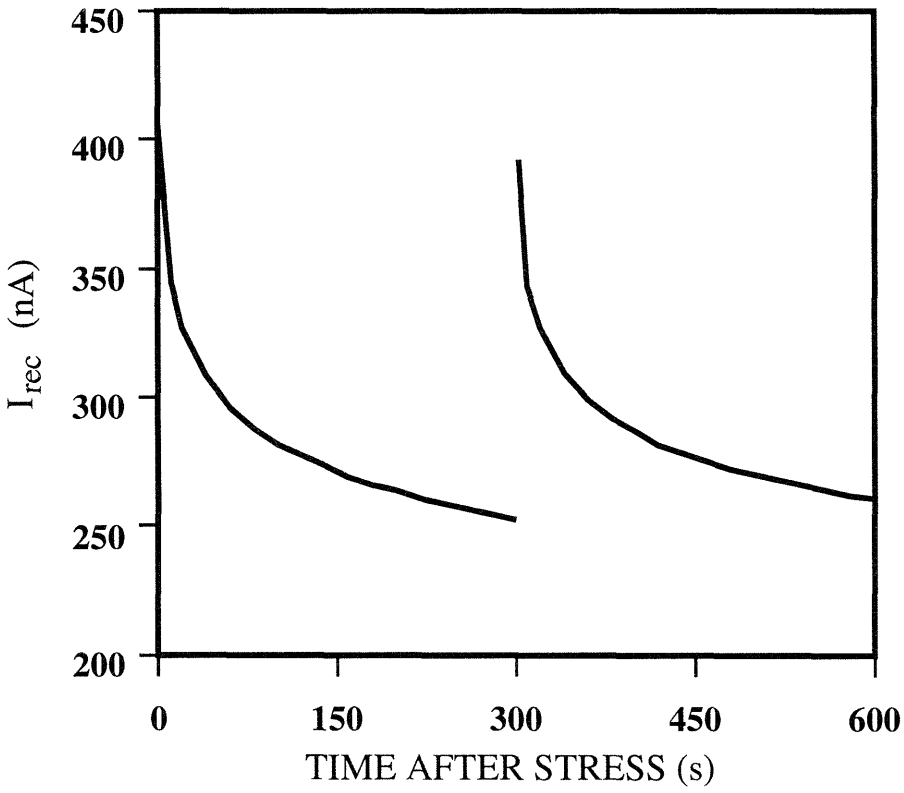


Fig. 12c. I_{rec} after a 5000 s stress.

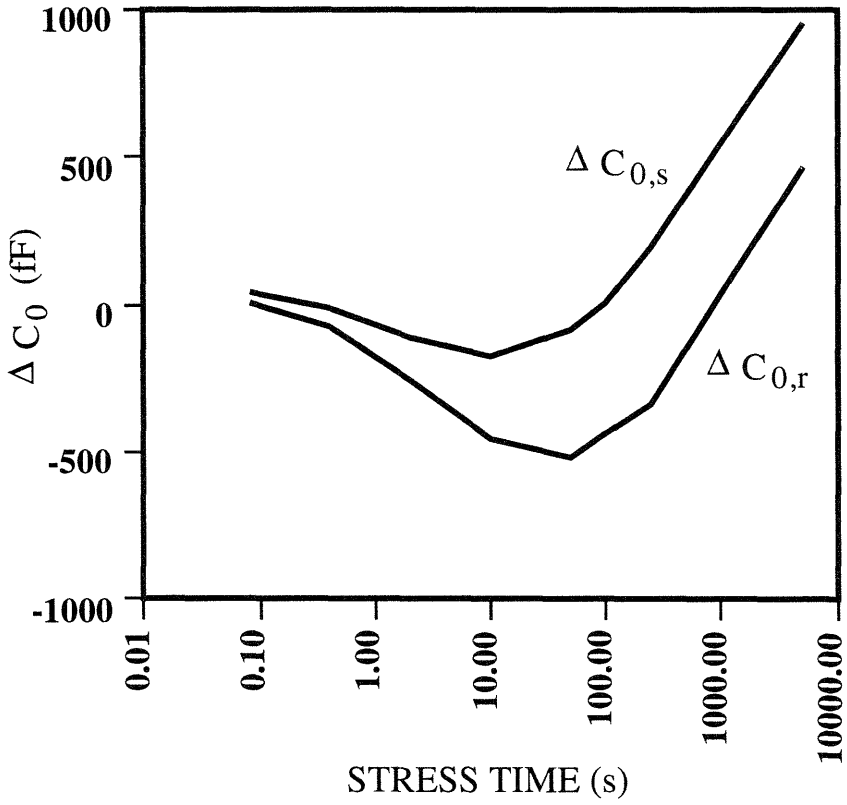


Fig. 13. $\Delta C_{0,s}$ and $\Delta C_{0,r}$ vs. stress time.

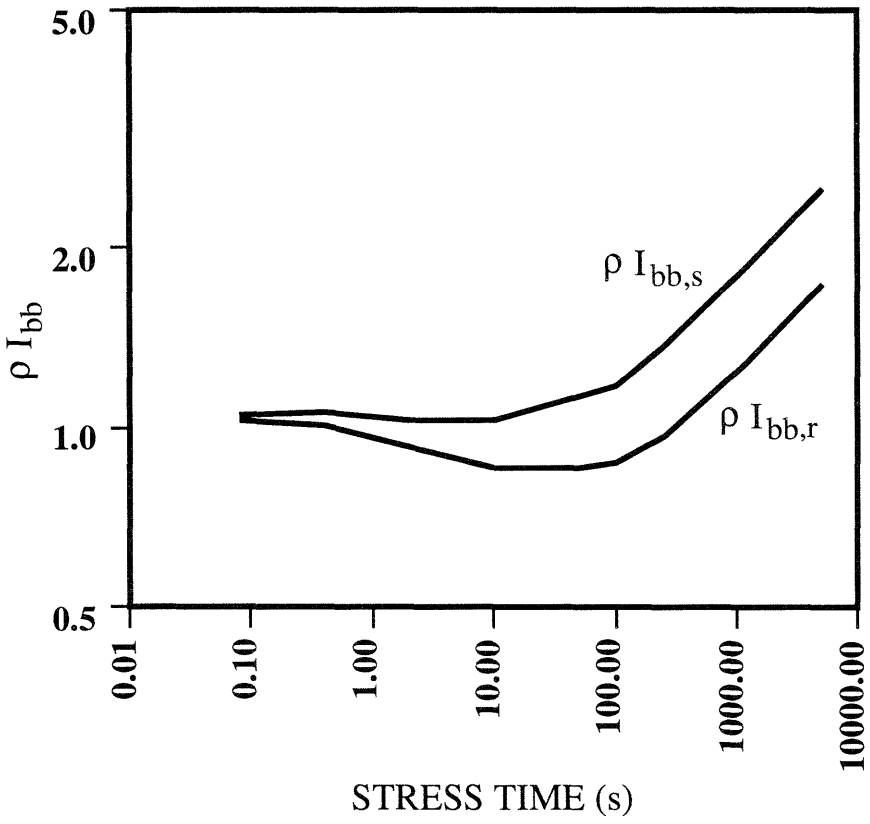


Fig. 14. $\rho I_{bb,s}$ and $\rho I_{bb,r}$ vs. stress time.

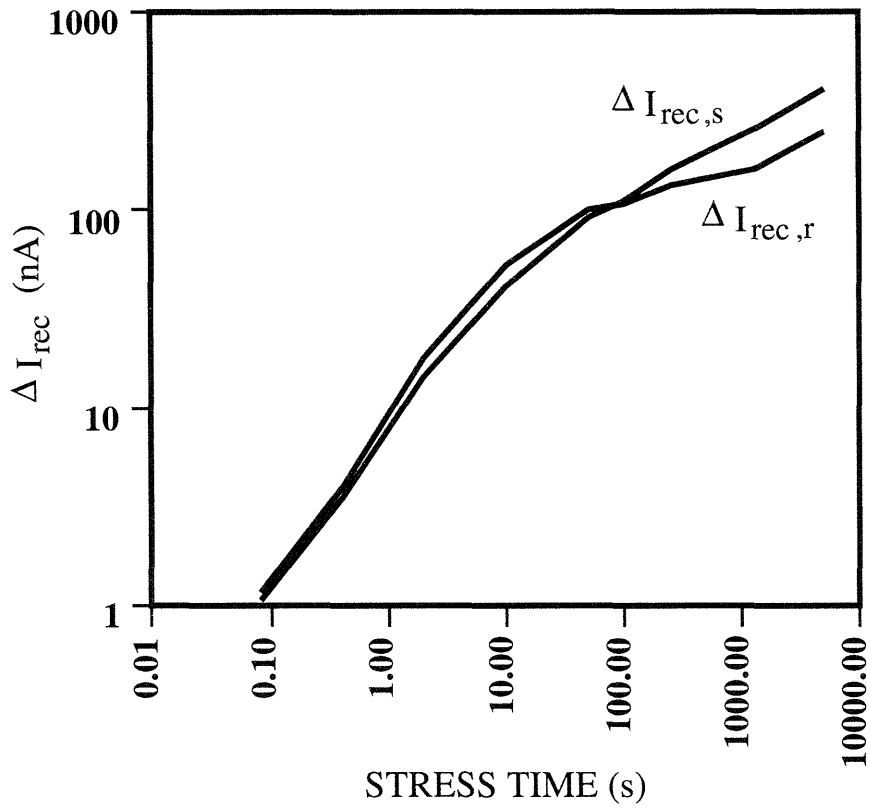


Fig. 15. $\Delta I_{rec,s}$ and $\Delta I_{rec,r}$ vs. stress time.

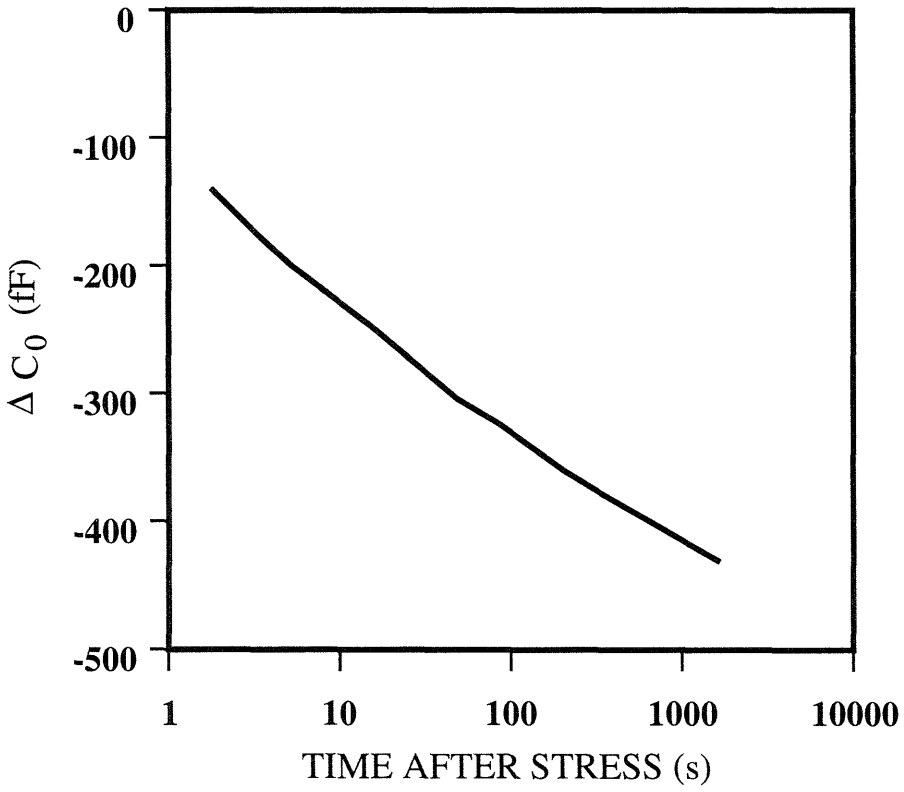


Fig. 16. ΔC_0 after a 6 s stress.

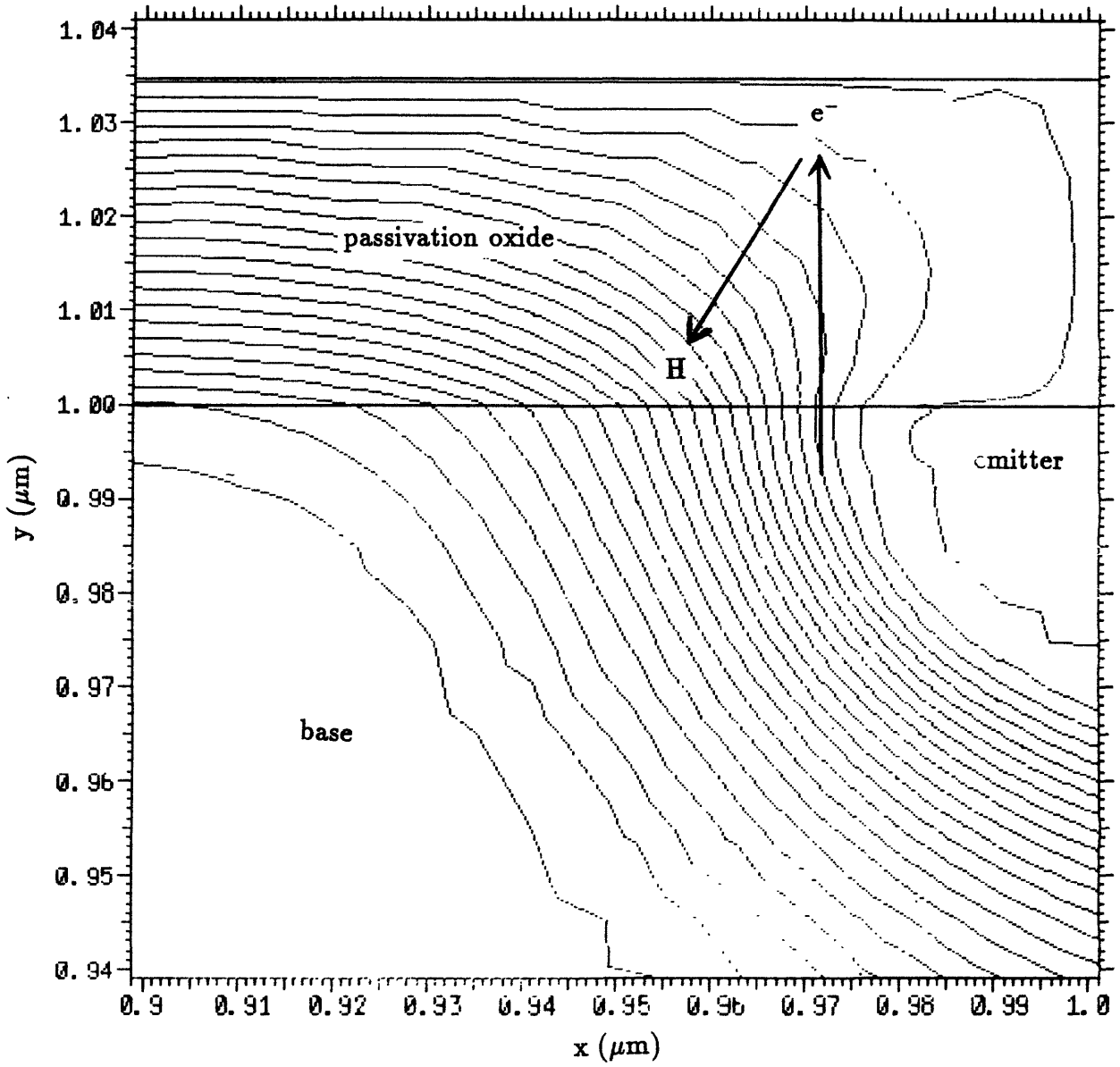


Fig. 17 Possible trajectories of a hot electron in the passivation oxide and the hydrogen it releases, shown with the potential contours during reverse-bias stress (see Fig. 5).

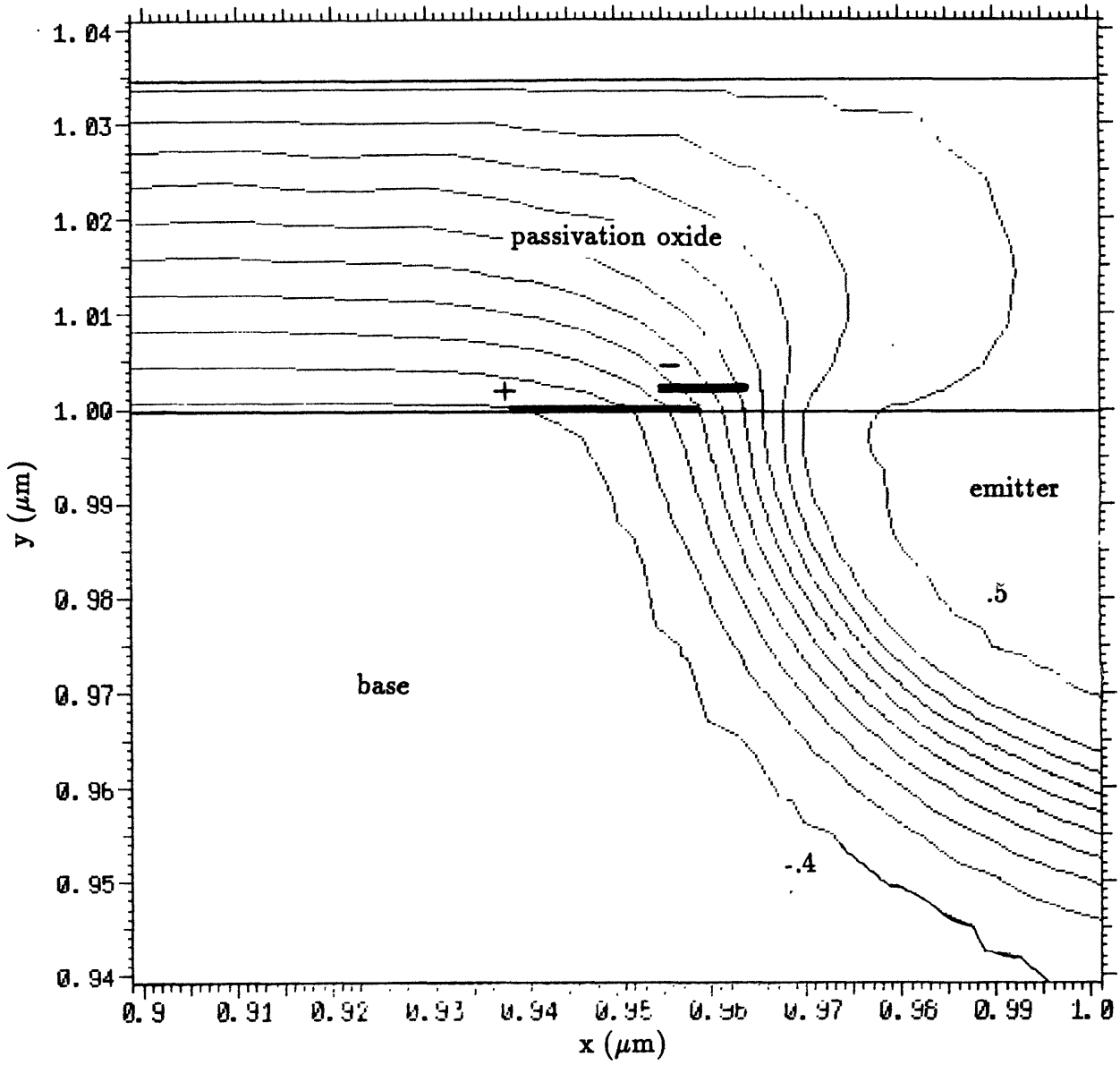


Fig. 18 Regions of positively and negatively charged defects in the passivation oxide of the experimental transistor, shown together with the potential contours at zero bias. The values of the potential on the contours run from -.4 V to .5 V in .1 V intervals.

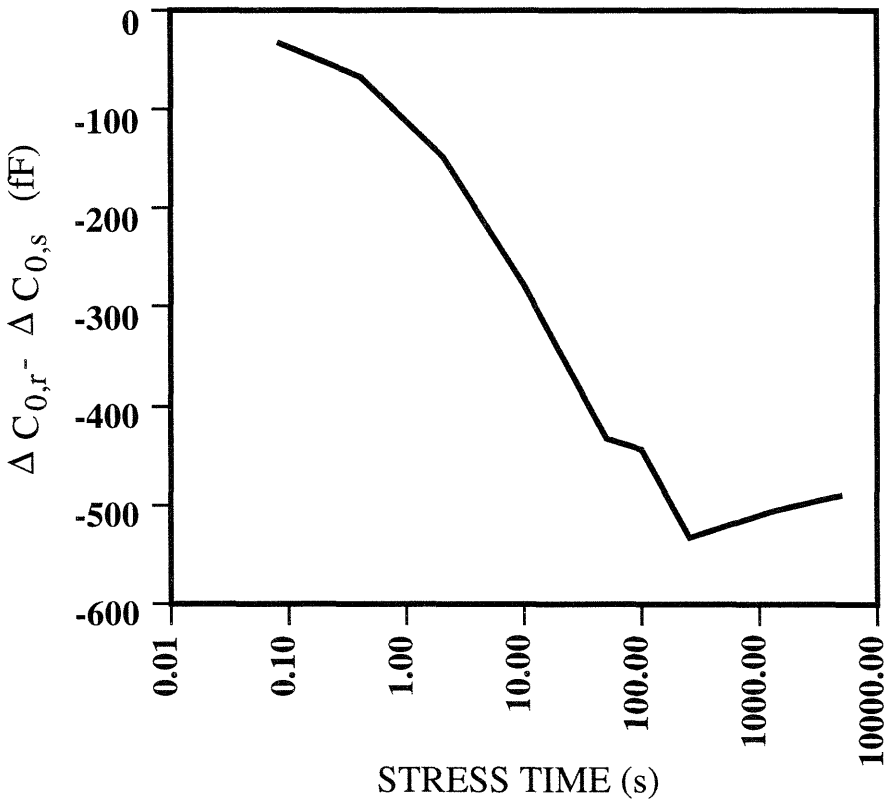


Fig. 19. $\Delta C_{0,r} - \Delta C_{0,s}$ vs. stress time.

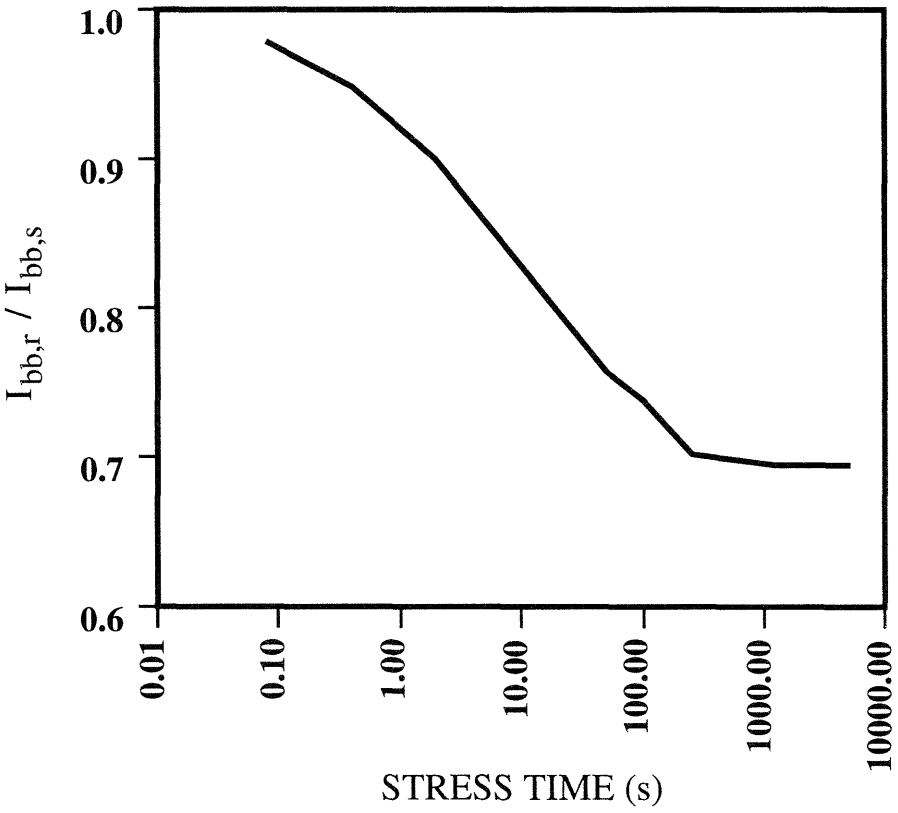
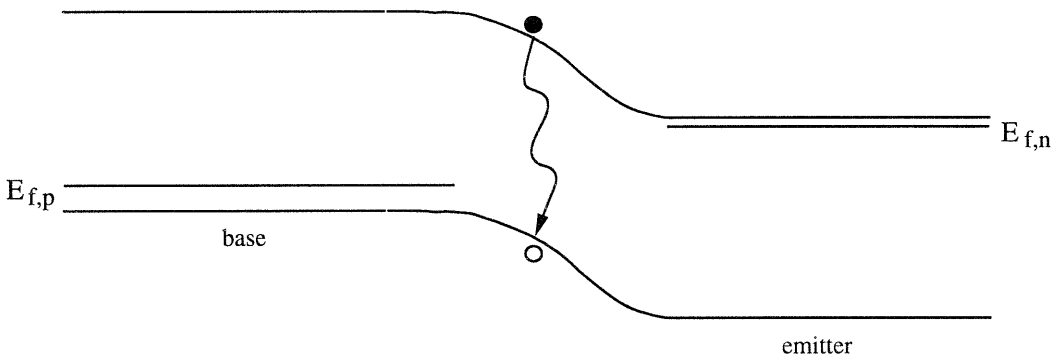
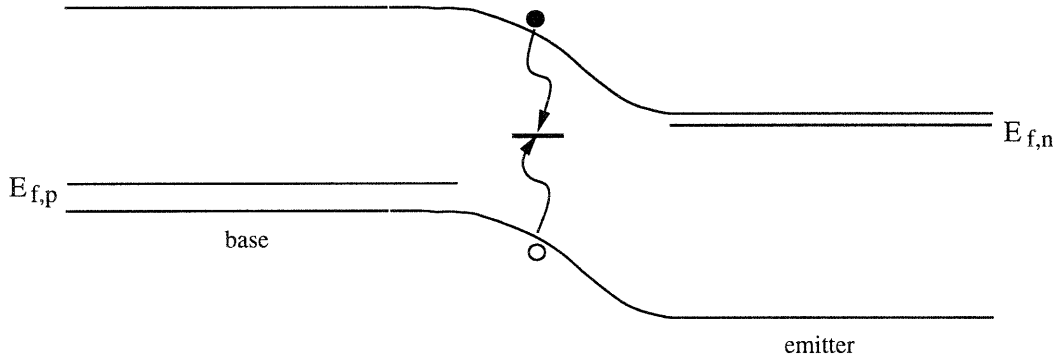


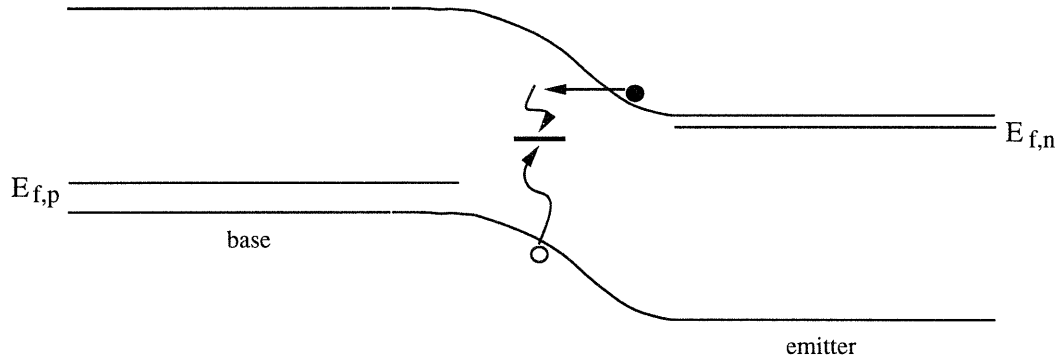
Fig. 20. $I_{bb,r}/I_{bb,s}$ vs. stress time.



a) Band-to-band recombination.



b) SRH recombination.



c) Field-enhanced recombination.

Fig. 21. Conduction band to valence band transit in forward bias.

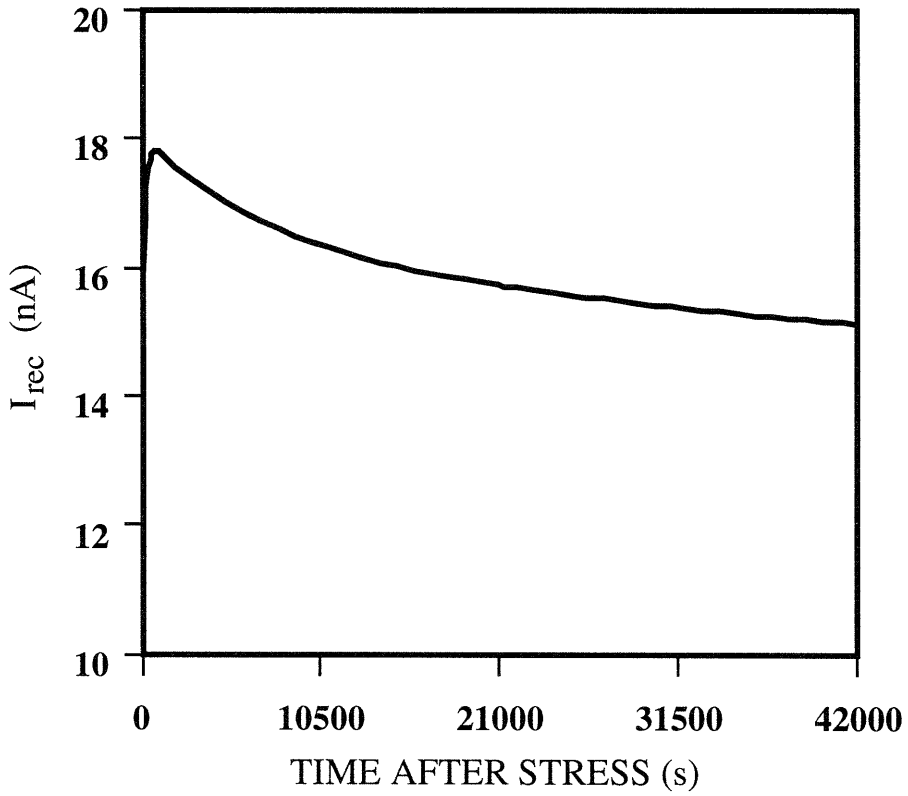


Fig. 22. I_{rec} after a 2 s stress.

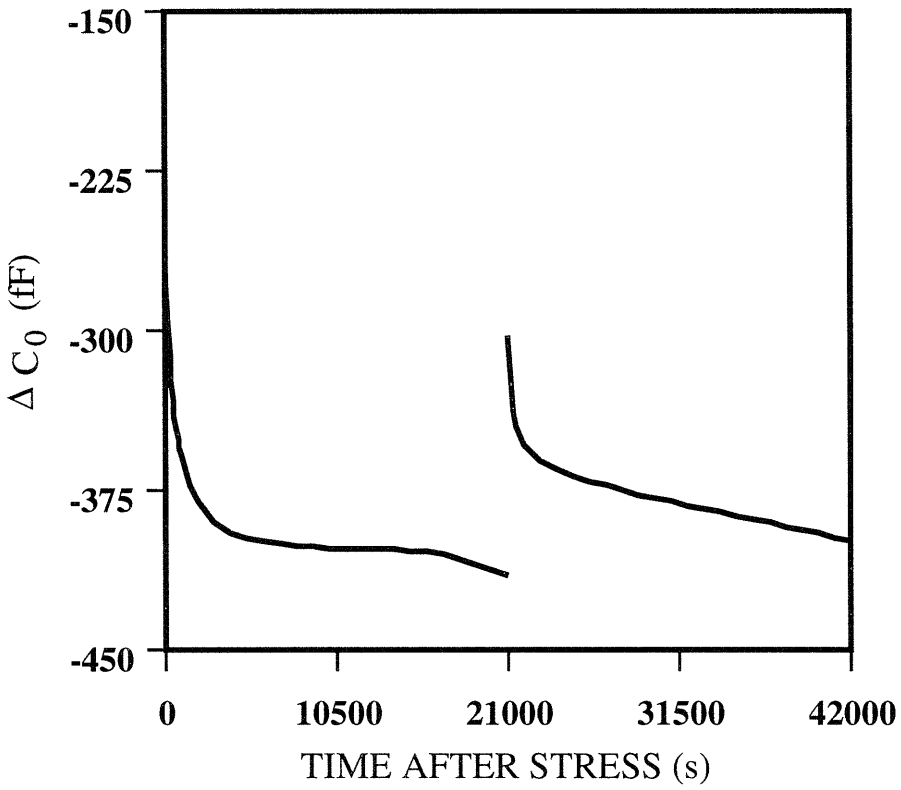


Fig. 23a. ΔC_0 after a 2 s stress. Relaxation is interrupted 21,000 s after stress by a forward-bias pulse of .8 V for 15 s.

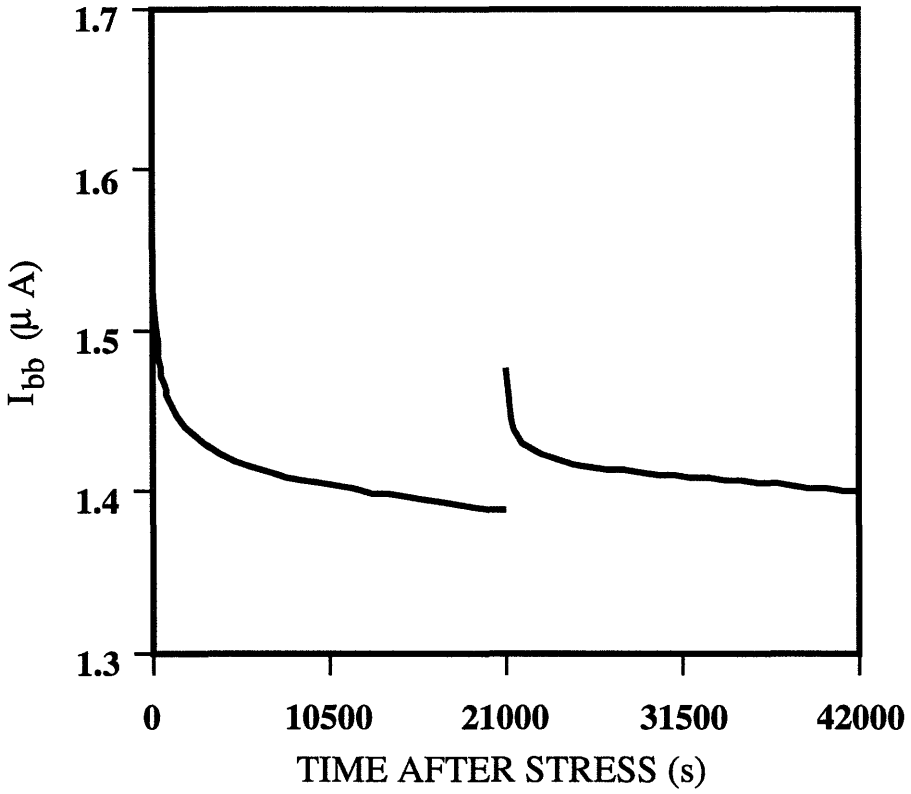


Fig. 23b. I_{bb} after a 2 s stress. Relaxation is interrupted 21,000 s after stress by a forward-bias pulse of .8 V for 15 s.

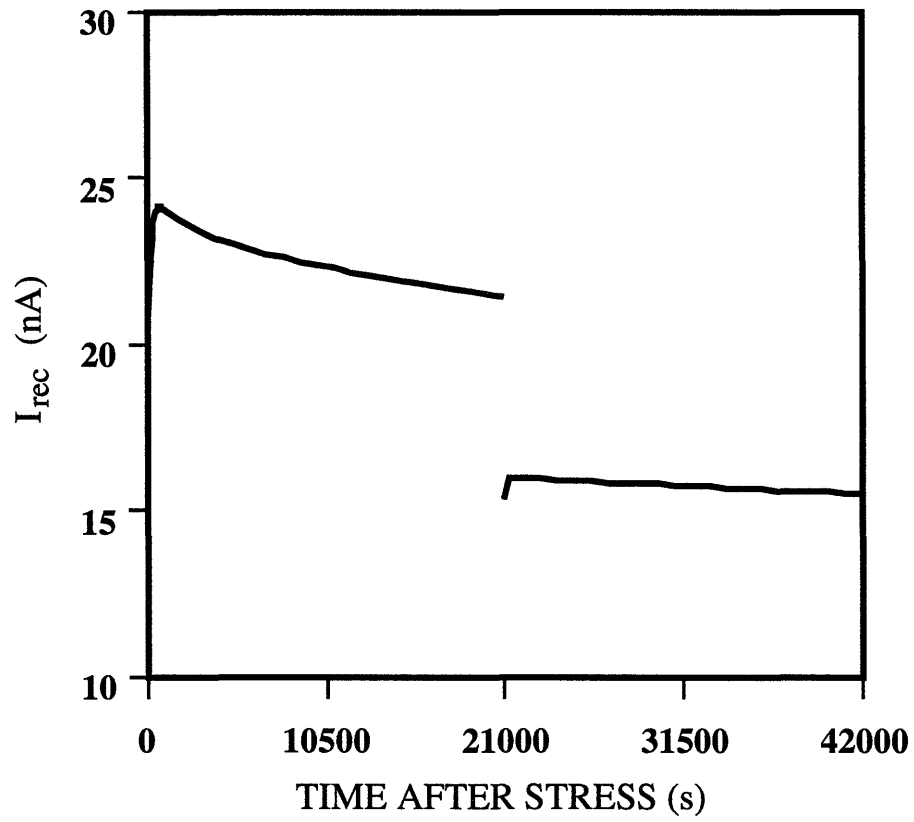


Fig. 23c. I_{rec} after a 2 s stress. Relaxation is interrupted 21,000 s after stress by a forward-bias pulse of .8 V for 15 s.

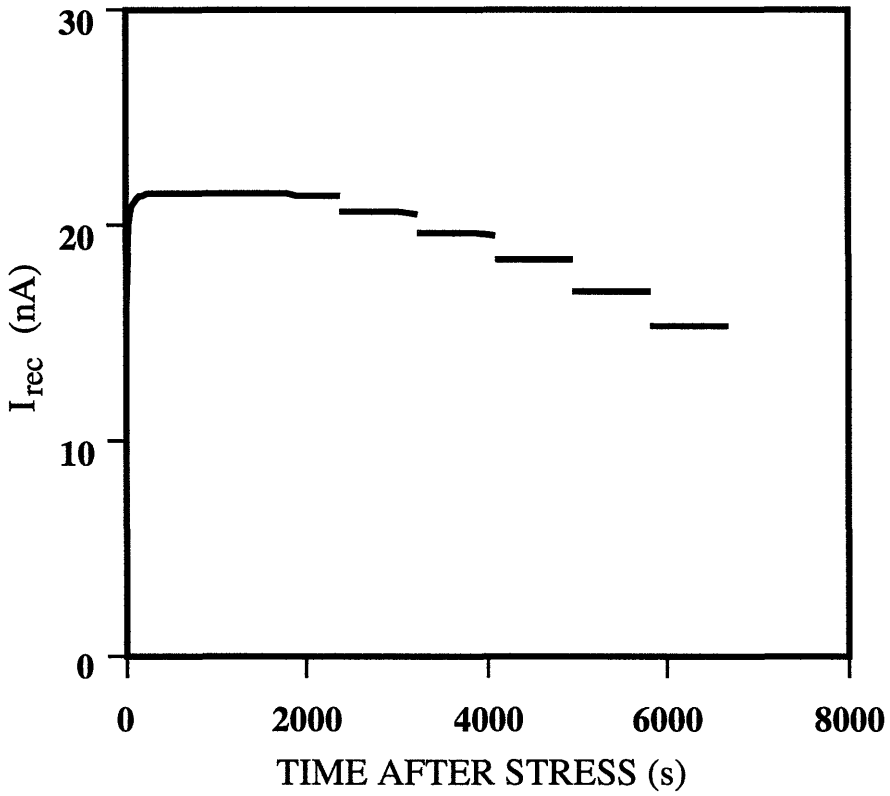


Fig. 24. I_{rec} after a 2 s stress. Relaxation is interrupted by five 1 s pulses of constant forward current. The currents range from 1 μA to 10 mA.

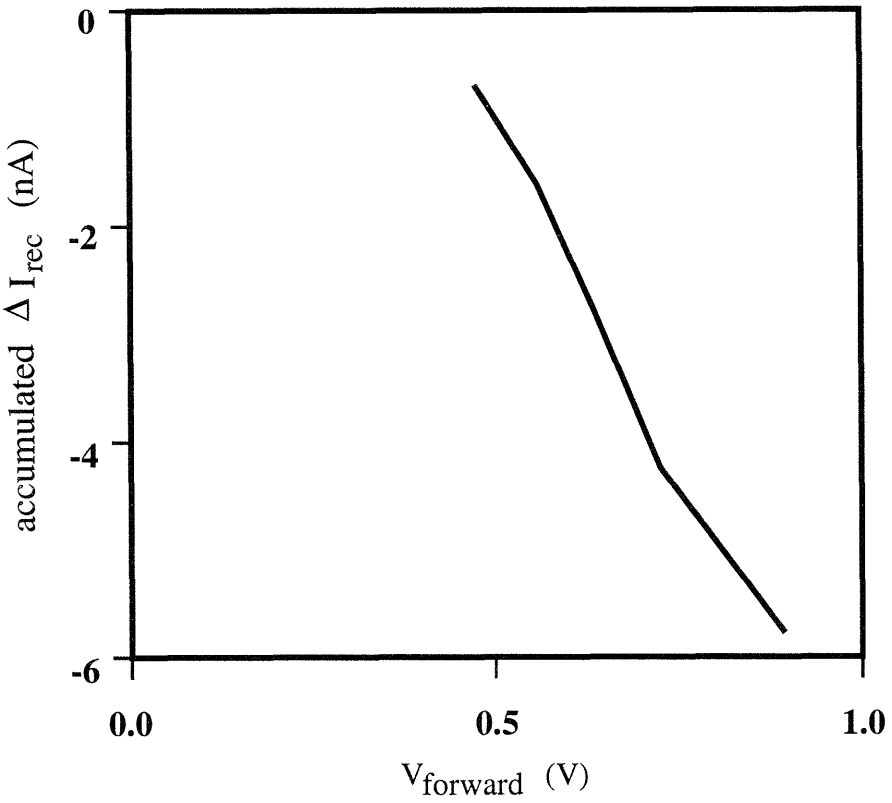


Fig. 25. Accumulated ΔI_{rec} vs. $V_{forward}$.

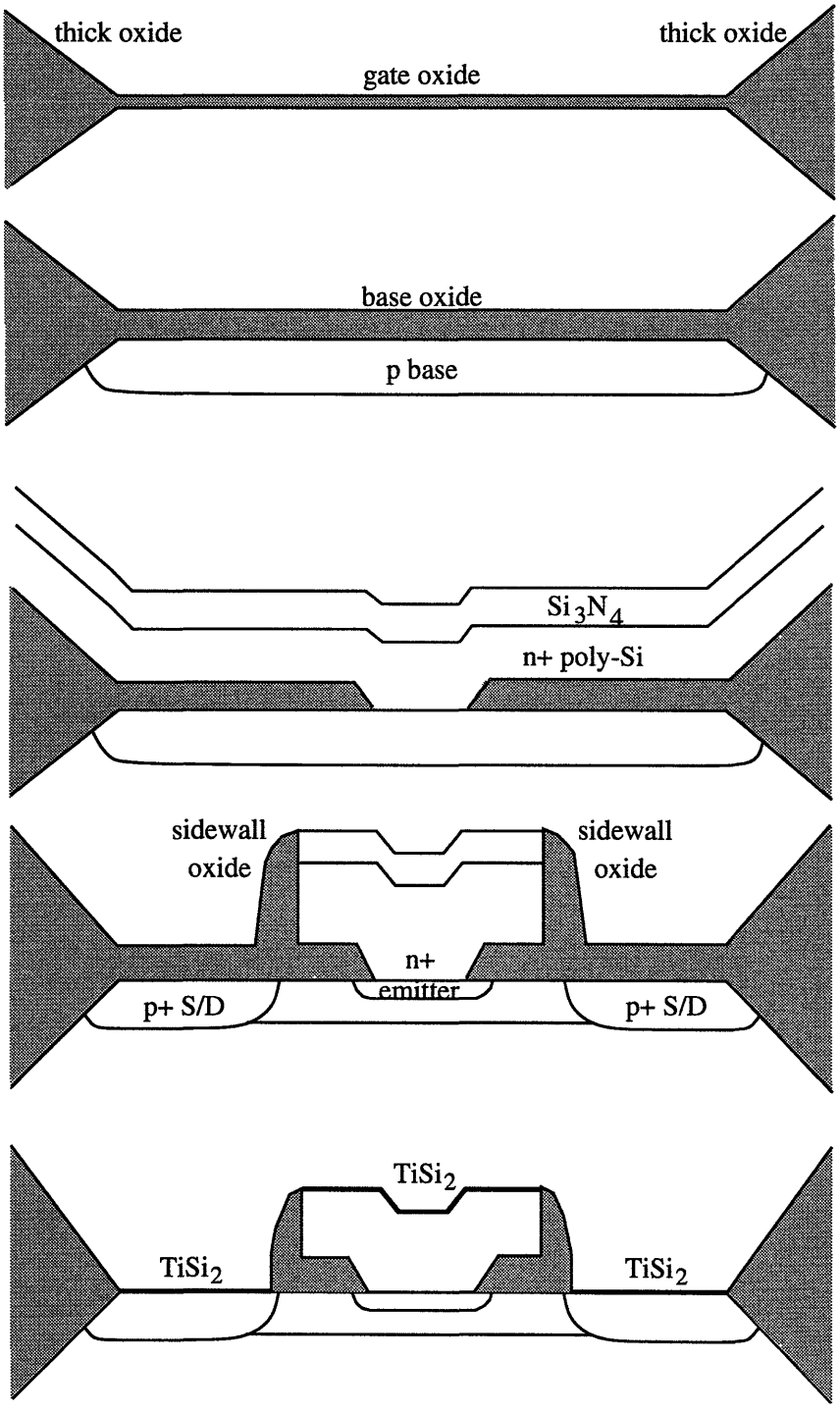


Fig. A1. Processing of the experimental transistor.

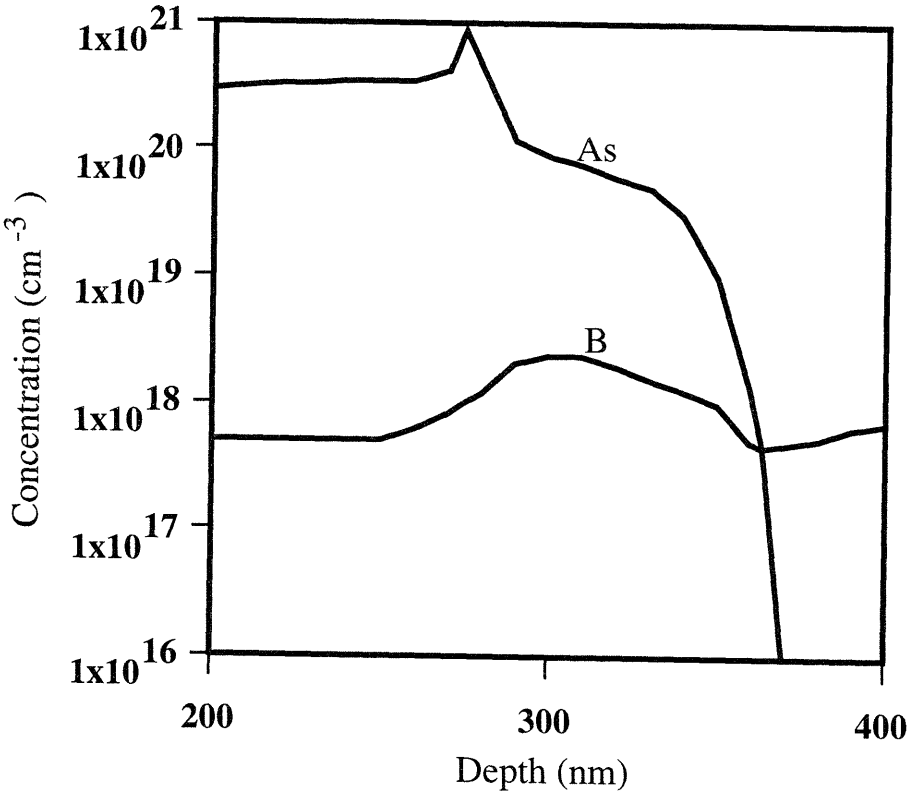


Fig. A2. Arsenic and boron profiles in the experimental transistor's emitter-base junction.

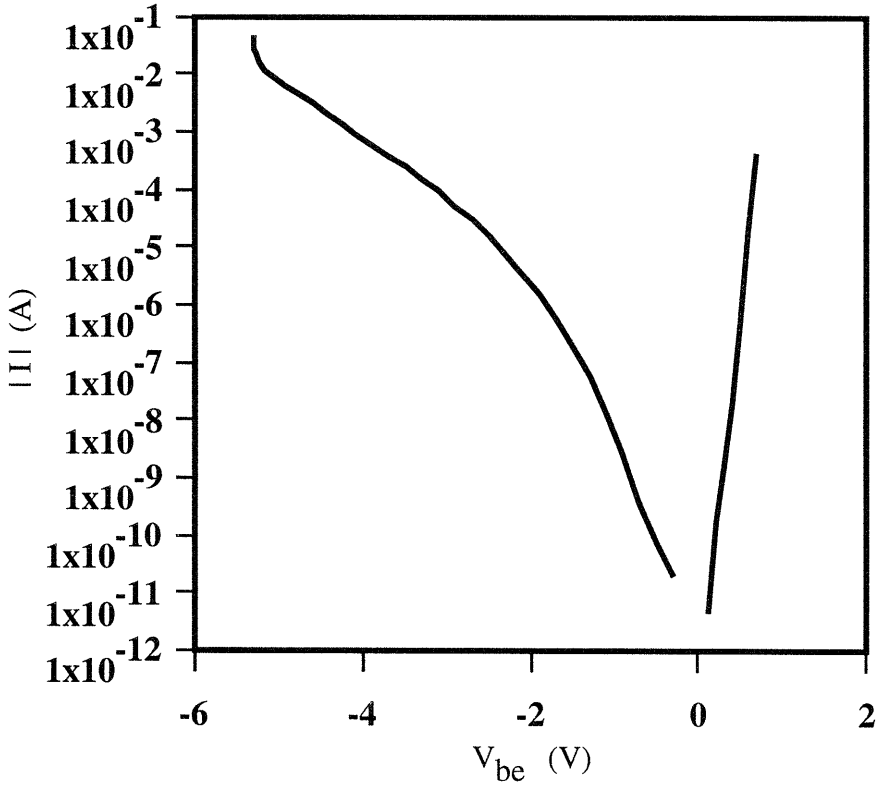


Fig. A3. I-V curve of the experimental emitter-base diode.

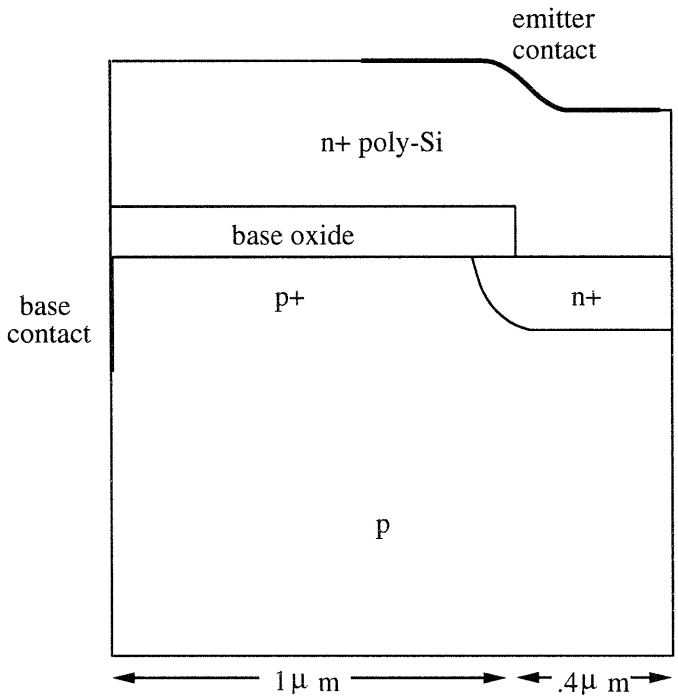


Fig. A4. Model of the emitter-base diode.

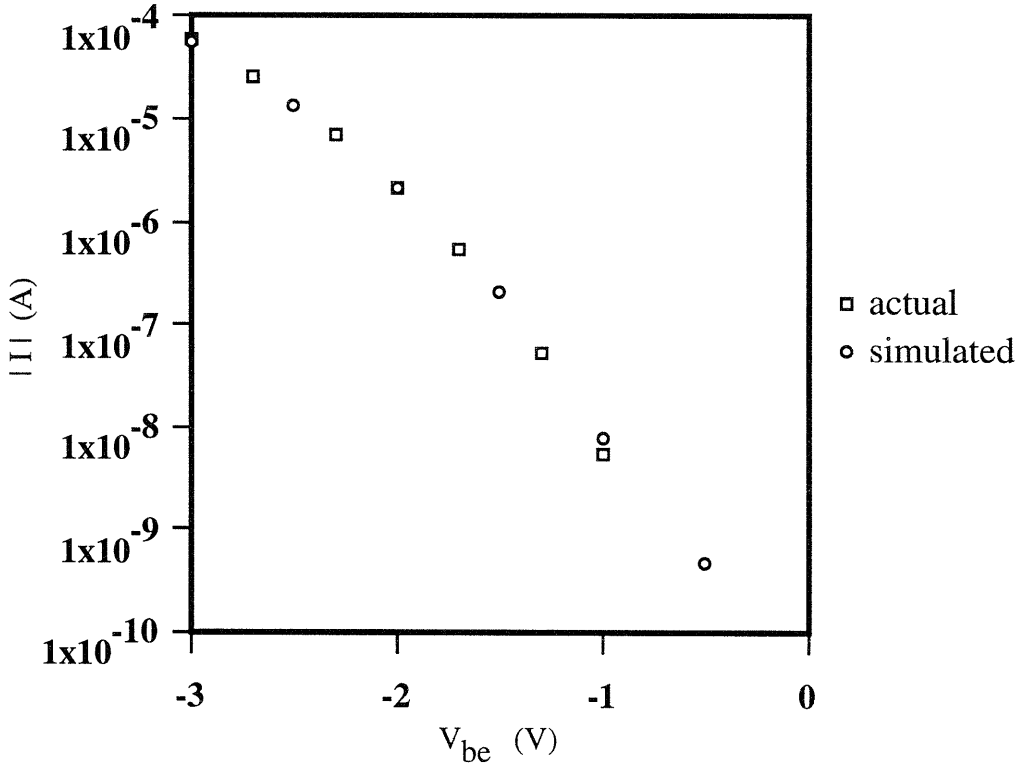


Fig. A5a. Actual and simulated I-V curves.

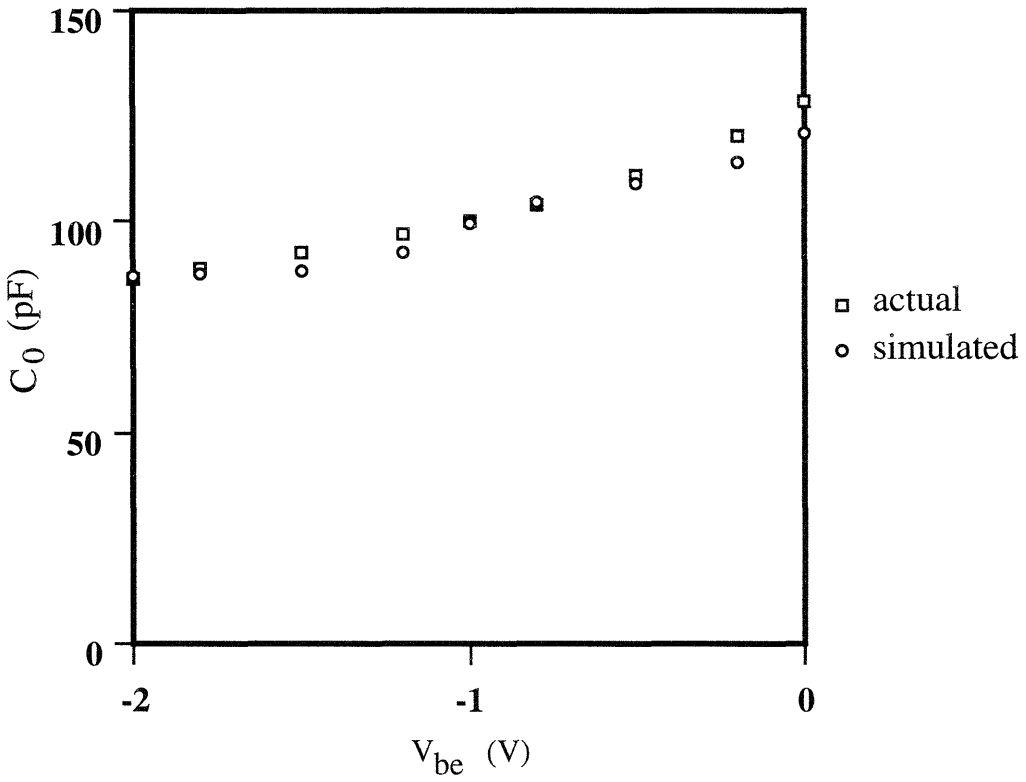


Fig. A5b. Actual and simulated C-V curves.

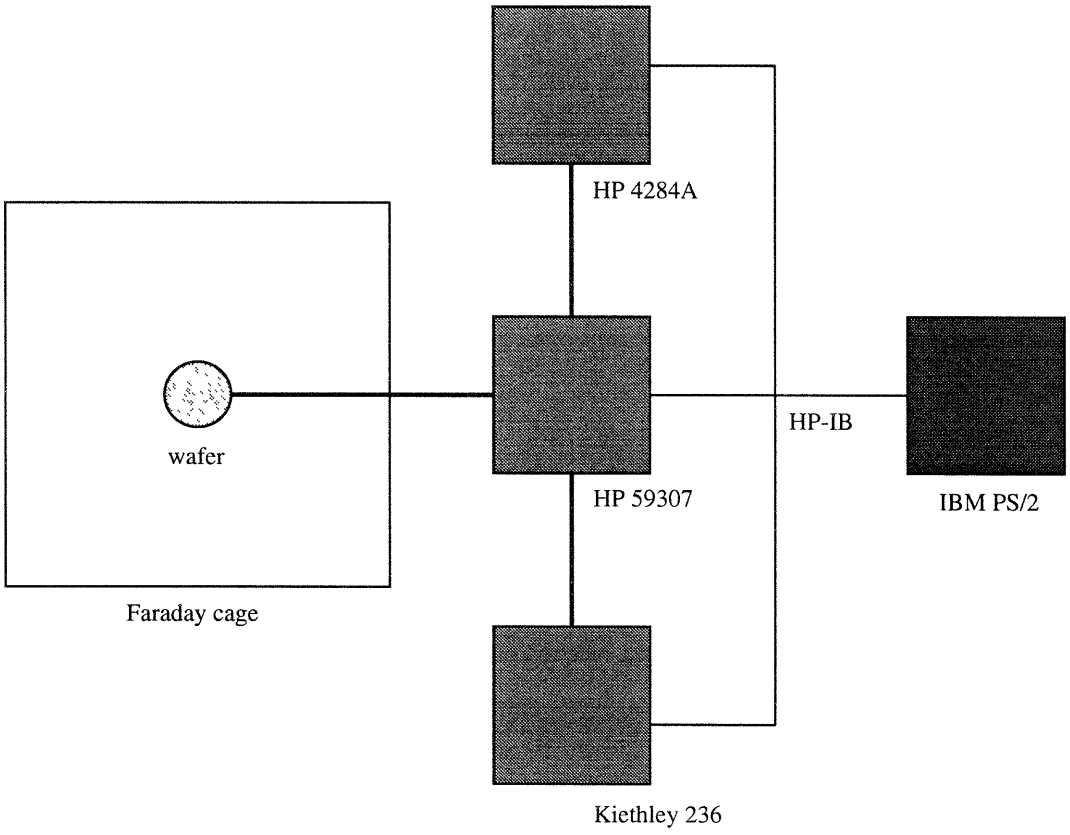


Fig. B1. Experimental setup.

References

[1] H. S. Momose *et al.*, *Proc. IEEE 1989 Bipolar Circ. and Tech. Mtg.*, p. 140, 1989.

[2] S. P. Joshi *et al.*, *IEEE Int. Electron Device Mtg. Technical Digest*, p. 182, 1987.

[3] Y. Niitsu *et al.*, *Proc. IEEE 1991 Reliability Phys. Symp.*, p. 193, 1991.

[4] L. Borucki *et. al*, *IBM Journal of Research and Development*, vol. 29, p. 263, 1985.

[5] E. M. Buturla *et al.*, *Proc. NASCODE 6 Conf.*, p. 291, 1985.

[6] E. H. Nicollian and J. R. Brews, *MOS Physics and Technology*, p. 497, John Wiliey and Sons, New York, 1982.

[7] J. F. Verwey, *Solid State Electronics*, vol. 14, p. 775, 1971.

[8] S. P. Joshi, *Proc. IEEE 1989 Bipolar Circ. and Tech. Mtg.*, p. 144, 1989.

[9] J. D. Burnett and C. Hu, *IEEE Trans. on Electron Devices*, vol. ED-35, p. 2238, 1988.

[10] S. A. Petersen and G. P. Li, *IEEE Int. Electron Device Mtg. Technical Digest*, p. 22, 1985.

[11] D. R. Collins, *Appl. Phys. Lett.*, vol. 13, p. 264, 1968.

[12] D. J. DiMaria and J.W. Stasiak, *J. Appl. Phys.*, vol. 65, p. 2342, 1989.

[13] C. Bulucea, *Solid-State Electronics*, vol. 18, p. 363, 1975.

[14] B. A. MacDonald, *IEEE Trans. on Electron Devices*, vol. ED-17, p. 871, 1970.

[15] A. Kapoor *et. al*, *Proc. IEEE 1991 Reliability Phys. Symp.*, p. 188, 1991.

[16] E. Hackbart and D. D. Tang, *IEEE Trans. on Electron Devices*, vol. ED-35, p. 2108, 1988.

[17] C. T. Wang, *Hot Carrier Design Considerations for MOS Devices and Circuits*, Van Nostrand Reinhold, New York, 1992.

[18] K. R. Hofmann *et al.*, *IEEE Trans. on Electron Devices*, vol. ED-32, p. 691, 1985.

[19] S. Tam *et al.*, *IEEE Trans. on Electron Devices*, vol. ED-31, p. 1116, 1984.

[20] P. E. Cottrell *et al.*, *IEEE Trans. on Electron Devices*, vol. ED-26, p. 520, 1979.

[21] E. Takeda *et al.*, *IEEE Trans. on Electron Devices*, vol. ED-29, pp. 611, 1979.

[22] J. D. Burnett and C. Hu, *Proc. IEEE 1990 Reliability Phys. Symp.*, p. 164, 1990.

[23] R. Bright and A. Reisman, *J. Electrochem. Soc.*, vol. 140, p. 2065, 1993.

[24] D. J. DiMaria *et al.*, *J. Appl. Phys.*, vol. 73, p. 3367, 1993.

[25] D. J. DiMaria *et al.*, *Appl. Phys. Lett.*, vol. 60, p. 2120, 1992.

[26] M Bourcerie *et al.*, *IEEE Trans. on Electron Devices*, vol. ED-37, p. 708, 1990.

[27] L. P. Trombetta *et al.*, *J. Appl. Phys.*, vol. 69, p. 2512, 1991.

[28] M. Fischetti, *J. Appl. Phys.*, vol. 57, p. 2860, 1985.

[29] G. A. M. Hurkx *et al.*, *IEEE Trans. on Electron Devices*, vol. ED-39, p. 331, 1992.

[30] A. Schenk, *Solid-State Electronics*, vol. 35, p. 1585, 1992.

[31] J. A. del Alamo and R. M. Swanson, *IEEE Electron Device Lett.*, vol. EDL-7, p. 629, 1986.

[32] G. A. M. Hurkx *et al.*, *IEEE Int. Electron Device Mtg. Technical Digest*, p. 307, 1989.

[33] C. J. Huang *et al.*, *Proc. IEEE 1991 Bipolar Circ. and Tech. Mtg.*, p. 170, 1991.

[34] M. Kerber, *J. Appl. Phys.*, vol. 74, p. 2125, 1993.

[35] D.V. Lang and L. C. Kimmerling, *Phys. Rev. Lett.*, vol. 33, p. 489, 1974.