Microsecond carrier lifetimes in strained silicon-germanium alloys grown by rapid thermal chemical vapor deposition

P. V. Schwartz and J. C. Sturm
Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544

(Received 11 June 1990; accepted for publication 20 August 1990)

We report generation lifetimes of over 1 μs in $Si_{0.82}Ge_{0.18}$ epitaxial strained layers grown by rapid thermal chemical vapor deposition on silicon substrates. By using a pulsed metal-oxide-semiconductor capacitor technique, we were able to probe the minority-carrier properties of a layer of $Si_{0.82}Ge_{0.18}$ sandwiched between two epitaxial layers of silicon. We also show that the band gap and the intrinsic carrier concentration are important when relating experimental results to the generation lifetime τ_g.

Silicon-germanium strained-layer alloys on silicon substrates have recently been applied to devices such as heterojunction bipolar transistors and *p-i-n* photodiodes.[1,2] Although these are minority-carrier devices, to date there have been no direct measurements of the minority-carrier properties of these films. In this letter we report the measurement of minority-carrier generation times of over 1 μs in strained $Si_{0.82}Ge_{0.18}$ layers grown on ⟨100⟩ silicon through the observation of the deep depletion recovery of metal-oxide-semiconductor capacitors.

The layers under study were grown in a susceptor-free, lamp-heated, chemical vapor deposition system similar to that described in Ref. 3. However, in this work, the sample temperature, in general, was not lowered to room temperature between the growth of adjacent layers. The starting wafers were 4 in., *p*-type, ⟨100⟩ CZ silicon wafers with resistivities of 10 Ω cm. Each substrate was chemically cleaned prior to loading and subsequently cleaned *in situ* by a hydrogen bake at 250 Torr at 1000 °C before the start of epitaxial growth. After the clean, a 1.5 μm *p*-type buffer layer, doped $\sim 3\times10^{17}$ cm^{-3}, was then grown at 1000 °C. A 30 nm $Si_{0.82}Ge_{0.18}$ strained layer was grown at 625 °C and the growth temperature was monitored *in situ* by optical transmission calibration.[4] The germanium composition was later confirmed by calibrated secondary-ion mass spectroscopy (SIMS), and the oxygen concentration of all layers was below 2×10^{18} cm^{-3}. Following the $Si_{0.82}Ge_{0.18}$ layer, a silicon cap layer of 30 nm thickness was grown at 850 °C for 30 s. Both the cap and the SiGe layer were also doped to $\sim 3\times10^{17}$ cm^{-3}. A sample device structure is shown in Fig. 1. Although not explicitly measured in this study, transmission electron microscopy (TEM) measurements on similar structures fabricated in our laboratory have shown that the samples are fully strained with a negligible number of misfit dislocations ($\geqslant 10\,\mu$m spacing). Control samples of all silicon layers were also grown at 700 °C (similar doping) for parallel fabrication and testing.

The gate oxides were formed by both plasma deposition and thermal oxidation. Low-temperature (400 °C) oxide plasma deposition (10 nm) followed by a 600 °C anneal in nitrogen was used to minimize the chance of strain relaxation in the silicon-germanium films and hence the formation of interface defects. A thermal oxide, grown at 800 °C in dry O_2, was also used as the gate insulator on some devices. The silicon gap layer was used so that the oxide/semiconductor interface was formed with silicon instead of silicon-germanium. (The SiO_2/Ge interface has been observed to have poor electrical properties.[5]) Since the thermal oxide consumes silicon, the cap layer of silicon in the final thermal oxide structures is estimated to be only 25 nm thick as shown in Fig. 1. The aluminum gate and guard ring were formed by thermal evaporation, followed by photolithography and etching. The guard rings, surrounding the gate, were to provide a means of limiting the active area of the device.

High-frequency (1 MHz) capacitance-voltage (*C-V*) curves of the capacitors showed well defined regions of accumulation, depletion, and inversion for both the silicon and the silicon-germanium buried layer samples. The dopings and oxide thicknesses extracted from the curves were consistent with the desired values discussed earlier. Straightforward calculations show that the depletion layer in inversion contains the buried layers of SiGe. Generation lifetimes were measured by observing the recovery of capacitance after pulsing the capacitor into deep depletion (Fig. 2) as in the Zerbst[6] technique. However, because of the multilayer structure of the sample, the details of the generation process must be examined before the lifetimes can be extracted.

Generation processes in bulk semiconductors can be modeled by the Shockley–Read–Hall equation:

$$G = \frac{v_{th} N_T (n_i^2 - np)}{1/\sigma_n (n + n_1) + 1/\sigma_p (p + p_1)}, \tag{1}$$

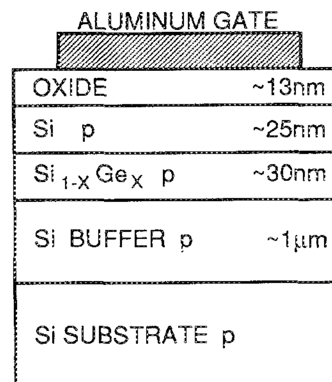

FIG. 1. Sample device structure for a buried layer silicon-germanium capacitor with a thermal oxide insulator. A voltage pulse applied between the gate and the substrate forces the capacitor into deep depletion with the depletion region residing primarily in the $Si_{1-x}Ge_x$ layer.

Downloaded 30 Jul 2002 to 128.112.49.151. Redistribution subject to AIP license or copyright, see http://ojps.aip.org/aplo/aplcr.jsp

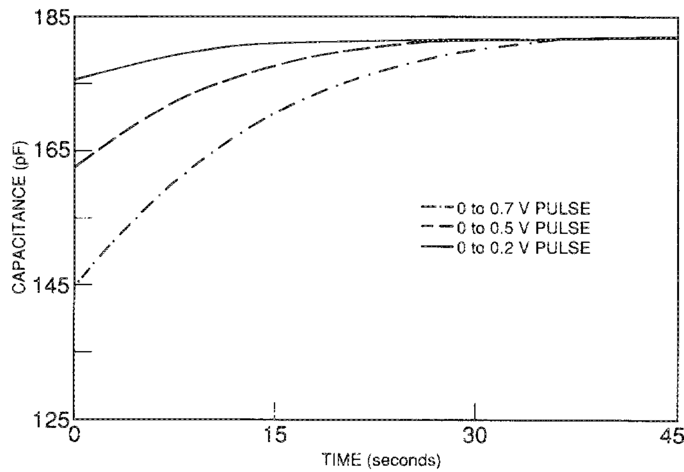

FIG. 2. Capacitance vs time plots for a buried layer $Si_{1-x}Ge_x$ capacitor. Time $t = 0$ represents the instant the voltage pulse is applied to the gate of the capacitor.

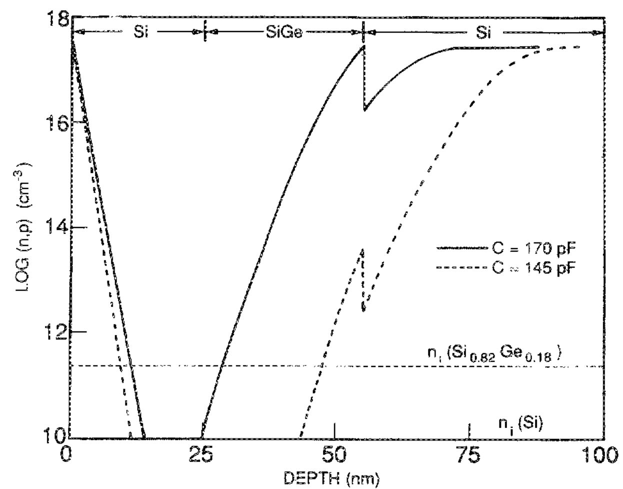

FIG. 3. Logarithm of the carrier profiles as a function of depth into the device for the buried layer $Si_{1-x}Ge_x$ capacitor at two points during its recovery from deep depletion to inversion. The discontinuity of the hole concentration reflects the discontinuity of the valence band at the $Si/Si_{1-x}Ge_x$ interface. The effective generation region is seen to fall primarily in the SiGe layer of the device.

where n_i is the intrinsic carrier concentration, n and p are the concentrations of the electrons and holes, respectively, v_{th} is the thermal velocity of the carriers, N_T is the trap density, σ_n and σ_p are the capture cross sections of the traps, and

$$n_1 = n_i e^{(E_t - E_i)/kT} \quad (2)$$

and

$$p_1 = n_i e^{(E_i - E_t)/kT}, \quad (3)$$

where E_T is the energy level of the trap and E_i the intrinsic Fermi level.

The generation process dominates only if the np product is less than n_i^2. If we assume $\sigma_n = \sigma_p = \sigma$ and that $E_T = E_i$ for simplicity, one finds

$$G = n_i^2/\tau_0(n + p + 2n_i), \quad (4)$$

where $\tau_0 = (\sigma N_T v_{th})^{-1}$. From this expression it is clear that generation is maximized only when both n and p are less than n_i. In this case, G is commonly written as n_i/τ_g, where $\tau_g = 2\tau_0$ and is called the generation lifetime. This expression cannot be applied to the entire depletion region, however because the carrier concentrations decrease with a finite slope when moving from neutral (high carrier concentration) regions to depletion regions. The situation is further complicated in our multilayer structures since the band gap of the $Si_{0.82}Ge_{0.18}$ layers is less than that of the silicon layers, resulting in a larger intrinsic carrier concentration. Assuming a band-gap reduction in the $Si_{0.82}Ge_{0.18}$ of 150 meV[7] compared to silicon, n_i in the silicon-germanium is about $n_i(\mathrm{Si})\exp(150/2kT) = 2.6\times10^{11}\ \mathrm{cm}^{-3}$ at room temperature. To illustrate this point, Fig. 3 shows the logarithm of the carrier profiles as a function of depth into the capacitor at two different times during its recovery from deep depletion to inversion for a gate-substrate voltage pulse from 0 to 0.7 V. This voltage range operates the capacitor in the inversion regime for the total range of its recovery. One can construct such a diagram in a straightforward manner from the capacitance which gives the depletion region width. (An electron quasi-Fermi level was calculated from the inversion layer density and a flat hole quasi-Fermi level in equilibrium with the substrate was assumed). The discontinuity in the hole concentration reflects the discontinuity of the valence band between the silicon and the silicon-germanium layers. As was stated before, maximum generation will occur when both n and p are less than the local n_i. During the entire recovery, this region is thus seen to sample within the $Si_{0.82}Ge_{0.18}$ layer, meaning that the generation within the SiGe region is not artificially suppressed because of high carrier concentrations.

A standard Zerbst plot (effectively plotting the generation rate versus generation volume) is shown in Fig. 4 for a buried silicon-germanium layer structure with a thermal oxide. However, before relating the slopes of this plot to the lifetime as is usually done, two points must be made. First, if one assumes as a worst case that all of the generation occurs in the SiGe films, the actual generation volume (SiGe only) is less than the total generation volume (includes SiGe and Si). Examination of Fig. 3 shows this correction to reduce

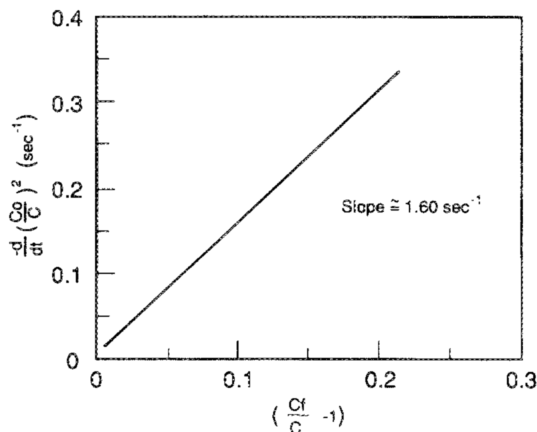

FIG. 4. Zerbst plot for the buried layer $Si_{1-x}Ge_x$ capacitor. The slope of the curve is related to the generation lifetime in the sample.

Downloaded 30 Jul 2002 to 128.112.49.151. Redistribution subject to AIP license or copyright, see http://ojps.aip.org/aplo/aplcr.jsp

TABLE I. Generation data for tested capacitors. The lifetimes computed for the samples containing $Si_{1-x}Ge_x$ layers used an adjusted value for the intrinsic carrier concentration of 2.6×10^{11} cm^{-3}. The silicon devices were pulsed from 0 to 3.0 V while the $Si_{0.82}Ge_{0.18}$ devices were pulsed from 0 to 0.7 V. The computed generation lifetime is represented as τ_g.

Sample	Doping ($\times10^{17}$ cm^{-3})	Oxide type	τ_g (μs)	Recovery time (s)
⟨100⟩ CZ silicon	0.005	thermal	9	325
⟨100⟩ CZ silicon	0.005	deposited	8	275
700 °C silicon	2	thermal	2	400
$Si_{0.82}Ge_{0.18}$	3	thermal	1	26
$Si_{0.82}Ge_{0.18}$	2	deposited	1	19

the extracted SiGe lifetime to be on the order of 50%. Second, the generation rate (and hence recovery time) are related by $G = n_i/\tau_g$. Since n_i depends on the band gap, one must use n_i for SiGe when calculating τ_g from the recovery time (slope of the Zerbst plot). A consequence of this is that for the same lifetime, generation rates in the SiGe layers are much greater than those in the silicon. Since generation is a thermally activated process of emission from traps to band edges, for the same trap density and cross section, one of course expects faster generation in a narrower gap material. The generation lifetime of this buried silicon-germanium layer sample is 1.45 μs.

The remainder of the lifetimes seen in Table I are found using an approximation to the Zerbst analysis as described by Schroder and Guldberg.[6] Several capacitors on each wafer were tested for lifetimes and the typical values are seen in Table I. (The generation lifetime found above and that found by the approximate method differ by only a factor of 1.5.) The recovery times (τ_r), shown in Table I, are defined as the time required for the capacitor to reach 90% of its final capacitance value after application of the voltage pulse. This parameter illustrates the difference in the recovery times between the buried silicon-germanium capacitors and the all-silicon capacitors. The low-temperature all-silicon capacitors (grown at 700 °C) that were fabricated along with the SiGe devices have total recovery times approximately 20 times longer than their silicon-germanium counterparts. If the generation were surface dominated, one would expect the total recovery time of the SiGe and the silicon capacitors (with similar doping) to be similar for large voltage pulses. The larger recovery time in the silicon demonstrates that the generation in the SiGe samples is most likely in the SiGe and not at the silicon/silicon dioxide interface. Comparing the lifetimes of the capacitors with deposited oxides and those with thermal oxides, one sees that the deposited oxides have lifetimes which are on the same order of magnitude. This is further evidence that the generation is not surface related. Although the silicon capacitors had a recovery time ~20 times longer than that of the SiGe capacitors, one finds similar lifetimes in the silicon and the SiGe films because of the different intrinsic carrier concentrations. This implies that the silicon and the silicon-germanium have similar trap densities and properties.

In conclusion, we have measured generation lifetimes in strained-layer $Si_{0.82}Ge_{0.18}$ grown by rapid thermal chemical vapor deposition of greater than 1 μs. In the experiment, the differences in the band gap, and consequently the local intrinsic carrier concentrations, are important when relating the generation rate to τ_g and when calculating the effective generation region. The long lifetimes demonstrate that one should be able to make low leakage junctions with this material as needed for heterojunction bipolar transistors and *p-i-n* photodiodes. The combined interest of Dr. A. Goodman of Office of Naval Research (No. N00014-88K-0396) and National Science Foundation (No. ECS-86157227) is appreciated. The assistance of C. Magee of Evans East, Inc. for SIMS is gratefully acknowledged.

[1] C. A. King, J. L. Hoyt, C. M. Gronet, J. F. Gibbons, M. P. Scott, and J. Turner, IEEE Electron Device Lett. **EDL-10**, 52 (1989).
[2] H. Tempkin, T. P. Pearsall, J. C. Bean, R. A. Logan, and S. Luryi, Appl. Phys. Lett. **48**, 963 (1986).
[3] J. F. Gibbons, C. M. Gronet, and K. E. Williams, Appl. Phys. Lett. **47**, 721 (1985).
[4] J. C. Sturm, P. V. Schwartz, and P. M. Garone, Appl. Phys. Lett. **56**, 961 (1990).
[5] T. O. Sedgwick, J. Appl. Phys. **39**, 5066 (1968).
[6] D. K. Schroder and J. Guldberg, Solid-State Electron. **14**, 1285 (1971).
[7] R. People, Phys. Rev. B **32**, 1405 (1985).

Downloaded 30 Jul 2002 to 128.112.49.151. Redistribution subject to AIP license or copyright, see http://ojps.aip.org/aplo/aplcr.jsp