Franz–Keldysh effect in the electric fields of macroscopic irregularities at a semiconductor surface

E. I. Gol'dman, A. G. Zhdan, and A. M. Sumaroka
Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 141120 Fryazino, Russia

(Submitted 12 May 1993)
Pis'ma Zh. Eksp. Teor. Fiz. **57**, No. 12, 783–787 (25 June 1993)

A phenomenological model for the generation of minority charge carriers at a semiconductor surface is constructed. This model successfully explains the exceedingly high rate of production of electron–hole pairs, which cannot be explained by other existing models. The generation kinetics predicted here agrees with experiments on silicon in the classical field-effect arrangement.

The activation energy for the spontaneous generation of electron–hole pairs in a semiconductor cannot be less than half the width of the band gap of the semiconductor in electric fields which are not too strong ($\leqslant 10^5$ V/cm).[1,2] For this reason, the time scale for the production of electron–hole pairs is $\tau \gg \tau_0 \exp(E_g/2T)$. For silicon ($E_g = 1.12$ eV), for example, this time must be more than tens of seconds near room temperature, even if the concentration of generation centers is high ($N \sim 10^{15}$ cm^{-3}). The rates of surface production of electron–hole pairs observed in Si at 300 K are three or four orders of magnitude higher than expected;[3] at lower temperatures, the discrepancy reaches ten orders of magnitude.[4]

In this letter we show that these discrepancies can be explained in terms of a thermally induced tunneling of charge carriers along the surface of a semiconductor in the electric fields of macroscopic irregularities. We examine the relaxation of the surface of a semiconductor, which we assume for definiteness is of n type, from a state of pronounced depletion to a state of pronounced inversion. In the latter case, the fluctuations of the surface potential are essentially screened by free holes. We assume that there are isolated "spots" at the surface in which the concentration of the internal positive charge is high, and we assume that the fields of these spots are not completely screened, to the point of equilibrium. The potential maxima corresponding to clusters of negative charges are "filled" primarily by holes.

Figure 1 shows the potential well near a spot of this sort, of radius a. The barrier to the escape of an electron from a potential well of depth U_0 created by the attraction toward the charges in a spot is lowered by the field of the depletion layer, $(2U_s/W) \times (1 - z/W)$, as a result of the Frenkel–Poole effect; here U_s is the surface potential, W is the width of the space-charge region, and z is the coordinate along the normal to the surface. The extent of this lowering, ΔU, is much larger than in the case of a singly charged center, since the number of net charges in a spot is $\gg 1$. Electron–hole pairs are produced through a thermally induced tunneling in the fringe field of the spot (as indicated by the arrows in Fig. 1a). The holes drain off along the surface, while the electrons remain in the well. They are subsequently emitted as a result of thermal energy across the barrier of height $U_0 - \Delta U$ into the interior of the semiconductor. In

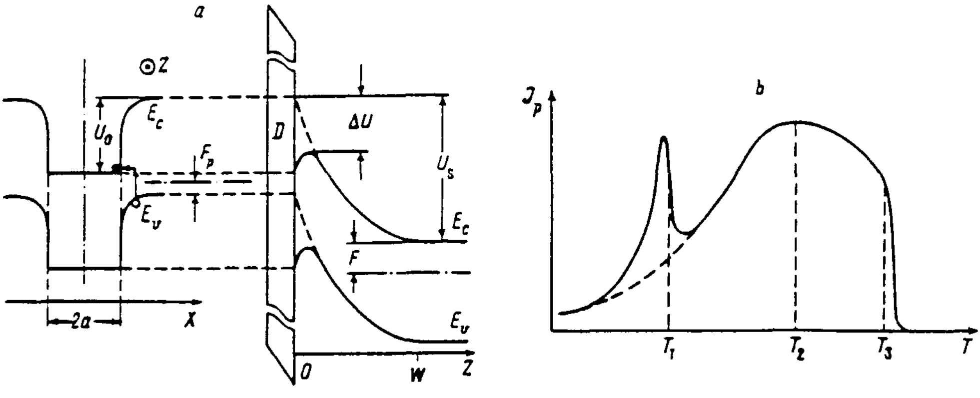

FIG. 1. a: Band diagram of a semiconductor near a charged spot. Left—In the plane of the surface; right—along the normal to the surface (E_c and E_v are the edges of the allowed band s of the semiconductor; D is an insulator). b: Qualitative temperature dependence of the rate of electron–hole-pair production in the field of a spot.

the approximation that the well is square along the surface,$\hbar^2 U_0/8mT^2a^2 \gg 1$ (m is the effective mass of an electron, and $\hbar$ is Planck's constant), the rate at which electron-hole pairs are generated at a spot, J_p, can be written

$$J_p = \nu_p \exp\left(-\frac{e_g U_0}{T}\right)\left[1-\exp\left(-\frac{E_g - F_p - F_n - U_0}{T}\right)\right]. \tag{1}$$

The electron flux through a barrier of height $U_0 - \Delta U$ is given by

$$J_n = \nu_n \exp\left(-\frac{U_0 - \Delta U + F_n}{T}\right)\left[1-\exp\left(-\frac{U_s + F - U_0 - F_n}{T}\right)\right]. \tag{2}$$

Here F_p, F_n, and F are the Fermi energies of holes at the surface and of electrons in the well and in the electrically neutral volume of the semiconductor, respectively; ν_p and ν_n are pre-exponential factors. The potential well serves as an effective generation center with some unusual properties: The activation energies for transitions of electrons into the conduction band, $U_0 - \Delta U$, and for transitions of holes into the valence band, $E_g - U_0$, depend on the filling of the well, on the hole concentration at the surface, and on the electric field of the depletion layer. Assuming that the number of electrons in the well is $\gg 1$, treating the spot as a charged metal disk, and using the known expression[5] for the potential of such a disk, we find $\Delta U = (4aU_sU_0/\pi W)^{1/2}$. Electron-hole pairs are produced under quasisteady conditions for the filling of the generation center. We equate J_p and J_n from (1) and (2).

$$(\nu_p/\nu_n)\exp[(U_0 - F_p - \Delta U)/T] \ll 1,$$

we find

$$J_p = (\nu_n \nu_p)^{1/2} \exp\left(-\frac{E_1}{T}\right), \tag{3}$$

where

$$E_1=\frac{E_g}{2}-\left[\frac{2aU_s}{\pi W}\left(E_g+T\ln\frac{\nu_n}{\nu_p}\right)\right]^{1/2}.$$

For high hole concentrations,

$$(\nu_p/\nu_n)\exp[(U_0-F_p-\Delta U)/T]\gg 1,$$

we find

$$J_p=\nu_n\exp\left(-\frac{E_2}{T}\right)\left[1-\exp\left(\frac{E_g-F_p-F-U_s}{T}\right)\right], \tag{4}$$

where

$$E_2=E_g-F_p-\left[\frac{4aU_s}{\pi W}(E_g-F_p)\right]^{1/2}.$$

Figure 1b is a sketch of the temperature dependence of the rate at which electron–hole pairs are produced. The results here are the same for isothermal relaxation (at a constant observation time at all T) and for thermally stimulated relaxation (T increases linearly with the time during the experiment). The part of the plot at $T<T_1$ is described by Eq. (3), which corresponds to the first stage of the generation, in which the concentration of the holes at the surface is still low. The temperature T_1 corresponds to the termination of the stage in which the holes being produced either localize at maxima of the surface potential or recombine with electrons trapped in surface states. In a narrow temperature interval at $T>T_1$, the hole concentration increases sharply, to the point that these holes reach a quasiequilibrium with electrons in the well. As a result, the activation energy for the generation rate increases to a value E_2; this change is seen as a peak near $T=T_1$ on the $J_p(T)$ curve. If holes did not localize or recombine at the surface, the transition to quasiequilibrium would have been smooth, and there would have been no maximum near this transition on the $J_p(T)$ curve (the dashed line in Fig. 1b). At $T>T_1$ the $J_p(T)$ dependence is described by (4), which corresponds to a quasiequilibrium of the electrons in the well and of the holes at the surface. The activation energy E_2 increases in the course of the relaxation, as the result of a decrease in the Fermi energy of the holes, a narrowing of the space-charge region, and an increase in the screening effect of the hole gas.[1)] The $J_p(T)$ curve has a relatively rounded maximum, which converts into a sharp drop near the temperature T_3, at which an equilibrium is reached. With increasing U_s, the activation energies E_1 and E_2 decrease, as do the temperatures T_1 and T_2. The temperature T_3 increases, as the result of an increase in the final equilibrium concentration of holes.

Let us compare the predicted kinetics of the production of electron–hole pairs with experimental data. Figure 2 shows the results of experiments carried out on n-Si (100) (type KÉF-1.0 silicon) in the standard arrangement for a time-varying field effect:[6] A state of strong nonequilibrium depletion is induced in the surface region of a sample cooled to 100 K. In the course of the thermally stimulated transition of the Si surface to an equilibrium state of pronounced inversion, three properties are mea-

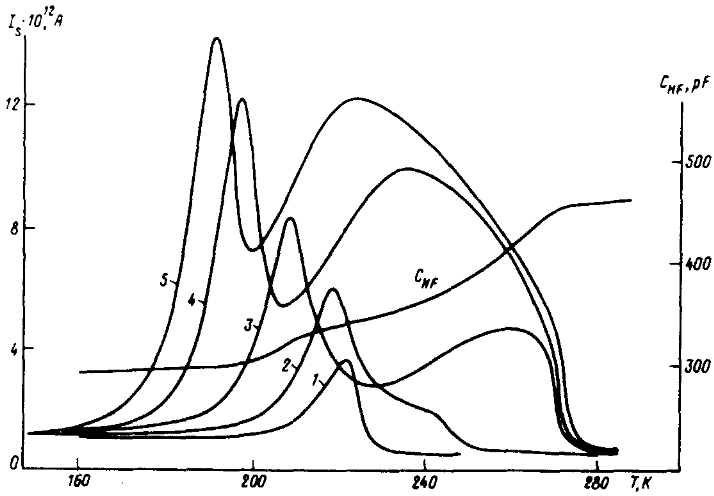

FIG. 2. Temperature dependence of the hole generation current at an n-Si-SiO_2 interface at various initial values of the space-charge field E_0 at a silicon surface. E_0, 10^4 V/cm: 1—2.8; 2—4.0; 3—4.5; 4–5.5; 5–8.5. The temperature dependence $C_{HF}(T)$, corresponding to curve 3, shows the typical behavior of the high-frequency capacitance in the course of hole generation. The area of the sample was 5.3×10^{-2} cm^2; the thickness of the insulator was 1750 Å.

sured simultaneously: the generation current $I(T)$, the small-signal (10-mV) capacitance of the system, and the high-frequency (1-MHz) capacitance of the system, $C_{HF}(T)$. Simultaneous measurements of pairs of $I_s(T)$ and $C_{HF}(T)$ curves made it possible to correct $I_s(T)$ for the displacement current due to relaxation of the width of the space-charge region on the basis of the local values of $C_{HF}(T)$.

Comparison of Figs. 1b and 2 shows that there is a complete qualitative agreement between the theoretical model discussed above and the experimental data. We find systematic explanations for the anomalously high rates of production of electron–hole pairs, the existence of two generation peaks (a narrow low-temperature peak and a broad main peak, which stretches up relatively high temperatures), the shift of these peaks toward lower temperatures with increasing initial band curvature U_{s0}, the sharp decrease in the generation current, and the slope change on the $C_{HF}(T)$ curves in the region in which the system goes to equilibrium. The rising branches of the low-temperature peaks of $I_s(T)$ are described by an Arrhenius law, with essentially the same pre-exponential factor for all U_{s0} values.

A plot of the corresponding activation energies E_1 versus $\mathscr{E}_0^{1/2}$ ($\mathscr{E}_0=2U_{s0}/qW$) is a straight line (Fig. 3), in agreement with (3). Point 1 is an exceptional case. The corresponding curve (1) in Fig. 2, with an activation energy $E_1\simeq E_g/2$, was found at a relatively small initial band curvature ($U_{s0}\simeq1$ eV), and it differs qualitatively from that shown in Fig. 1b. Apparently, one of the "classical" channels for the generation of electron–hole pairs becomes predominant upon the conversion to weak inversion. The linear $E_1(\mathscr{E}_0^{1/2})$ plot intercepts the energy axis at 0.55 eV, which is very close to $E_g/2$. The slope of this plot is determined by the spot

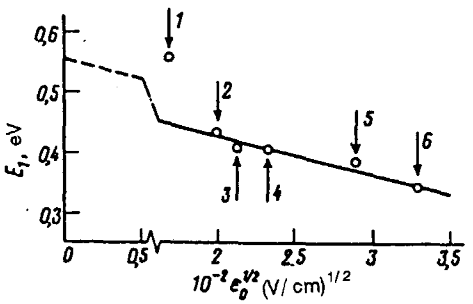

FIG. 3. Activation energy for the rising branches of the low-temperature peaks, E_1, versus the electric field of the space charge near the silicon surface, E_0. The numbers on the arrows correspond to the curve labels in Fig. 2. The $I_s(T)$ dependence corresponding to point 6, with $E_0=1\times10^5$ V/cm, is not shown in Fig. 2.

radius $a=86$ Å. Working from the experimental value of the pre-exponential factor, assuming $v_n=v_p\sim a^2V_tN_c$, where V_t is the thermal velocity of the electrons, and N_c is the effective density of states in the Si conduction band, we estimate the concentration of charged spots to be $\sim 10^5$ cm^2. In a state of pronounced inversion we would have $U_0\approx E_g$, so the number of positive charges in a spot would be at least 25, and their concentration higher than 10^{13} cm^{-2}.

In summary, the nature of the anomalously high generation of electron–hole pairs at a semiconductor surface can be explained as a manifestation of the Franz-Keldysh effect in the electric fields of macroscopic irregularities. Estimates show that even a vanishingly small number of such irregularities, which would be difficult to detect by direct methods, would have a fundamental effect on the rate of production of electron–hole pairs.

[1)]This circumstance was ignored in Eq. (4).

[1]J. G. Simmons and H. A. Mar, Phys. Rev. **B8**, 3865 (1973).
[2]H. A. Mar and J. G. Simmons, Phys. Rev. **B11**, 775 (1975).
[3]S. M. Sze, FPP **1**, 444 (1984); S. M. Sze, *Physics of Semiconductor Devices* (New York, 1981).
[4]D. Vuillaum, M. Lannoo, J. C. Bourgoin, and E. Resencher, J. Appl. Phys. **66**, 5920 (1989).
[5]L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1984).
[6]A. V. Rzhanov, *Electronic Processes at Semiconductor Surfaces* (Nauka, Moscow, 1971), p. 244.

Translated by D. Parsons