Intermediate state of a controllable four-layer $p-n-p-n$ structure

N. Z. Vagidov, Z. S. Gribnikov, A. N. Korshak, and V. V. Mitin*

Institute of Semiconductor Physics, National Academy of Sciences of Ukraine 252028 Kiev, Ukraine
(Submitted November 25, 1994; accepted December 6, 1994)
Fiz. Tekh. Poluprovodn. **29**, 1958–1975 (November 1995)

The intermediate state considered here is the state of incomplete blocking of the structure, for which the gate current is inadequate to completely switch off the anode current but substantially compresses the conductive region, increasing the current density in it. This incomplete-shutoff regime can be useful for controlling light-emitting structures and can make it possible to adjust the size of the luminous region, its position, its radiance, the relationship between the various recombination channels, etc. A steady-state theory of the potential distribution and the current densities in the incomplete-shutoff regime of a four-layer structure has been developed for various values of the total anode and gate currents. The current limits have been found for the existence of heterogeneous states, including the conducting and "blocked" regions. The effect of the finite conductivity of the controlled (high-resistance) base has been taken into account. Computations have been carried out for the very simple case of a linear recombination law in each base and at small injection levels at least in the controlling (low-resistance) base. © *1995 American Institute of Physics.*

1. The object of study in this paper is the four-layer semiconductor $p^+-n-p-n^+$ structure shown in Fig. 1. A segment of the structure having the shape of a long uniform strip with a width of $2l$ is investigated (in an actual situation, this is the width of the cathode strip). The outer p^+ and n^+ regions (at the bottom and top in Fig. 1) are assumed to be so strongly doped that they unconditionally ensure that the injection coefficients equal unity for the majority carriers in the middle base regions (the inner regions of the circuit shown in Fig. 1). Nonequilibrium phenomena in the outer p^+ and n^+ regions are therefore disregarded. Such phenomena are considered in the inner p and n regions—the bases of the structure. One of these bases— the controlling p base—is equipped with side electrodes at $y=\pm l$ (Fig. 1). These side electrodes, which are ohmic p^+ contacts to the controlling p base, are called gates (following American usage).

We shall consider the open-structure regime, in which all three $p-n$ junctions (numbered *1*, *2*, and *3* in Fig. 1) are biased in the forward direction, and a large forward current flows through them. If the gate circuits are interrupted and there is no current in either gate, $\mathscr{J}_{g1}=\mathscr{J}_{g2}=0$, a homogeneous anode current $j_a(0)=\mathscr{J}_a/2l$ flows through the entire width of the device; $\mathscr{J}_a$ is assumed to be given. We shall supply identical potentials $\varphi_{g1}=\varphi_{g2}=\varphi_g$ at the gates, negative with respect to the cathode, which provide the gate currents $\mathscr{J}_g$. These biases tend to block not only by cathode $p-n$ junction *1*, but also the middle $p-n$ junction *2*. Such blocking can be considered approximately homogeneous, provided the following condition is satisfied:

$$\delta\varphi_g=\frac{\mathscr{J}_g l}{\sigma_p}<\frac{T}{e}, \tag{1}$$

where σ_p is the longitudinal conductivity of the controlling p base, and T is the temperature in energy units. In this case, the structure under consideration behaves like a homogeneous device which is characterized by a single current density j_a, approximately identical in the entire interval $(-l, l)$. As $\mathscr{J}_g$ increases, so that inequality (1) breaks down and is replaced by the inequality with the opposite sense, the structure of the anode current density becomes essentially heterogeneous: A central conductive region $|y|<x_c$ appears. In this region the current density increases compared to the initial value,

$$j_a(\mathscr{J}_g)=\frac{\mathscr{J}_g}{2x_c(\mathscr{J}_g)}>j_a(0). \tag{2}$$

A peripheral blocked region $x_c<|y|<l$ also appears. The width of this region increases with increasing $\mathscr{J}_g$. There is a certain critical value of $\mathscr{J}_g$, which can be calculated and which equals $\mathscr{J}_M(\mathscr{J}_a)$; when this value is exceeded, the device is completely blocked, and the conductive region disappears. In the range

$$\frac{\sigma_p T}{el}<\mathscr{J}_g<\mathscr{J}_M(\mathscr{J}_a) \tag{3}$$

we are dealing with a certain heterogeneous intermediate state of the structure, which is partly blocked and partly open, so that the higher the current $\mathscr{J}_g$ (for a given $\mathscr{J}_g$) and the larger the area of the blocked part of the structure, the larger is the current density in the open part. This incomplete-shutoff regime (ISR) is the main subject of consideration here.

We should point out that the ISR considered here is of no great interest in power thyristors, since decreasing the conductive region means increasing the current density j_a in it and, consequently, increasing the voltage φ on the device, while the volume in which the evolved power is dissipated simultaneously decreases. Therefore, the only transition process of interest in power devices is that of complete blocking, which switches off the anode current in thyristors blocked from the base [i.e., in gate turn-off (GTO) thyristors].

 1063-7826/95/111021-09$10.00

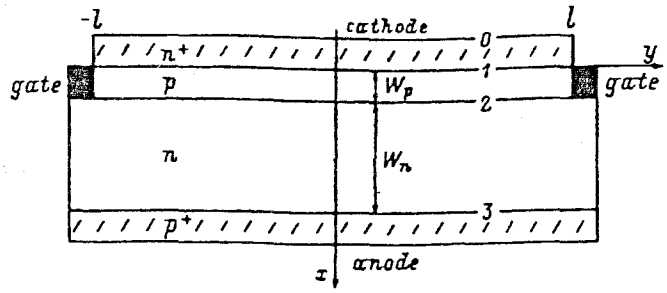

FIG. 1. Cross section of four-layer $p^+-n-p-n^+$ structure.

The ISR, briefly considered by Gribnikov and Rothwarf in Ref. 1, which is devoted to the theory of the GTO thyristor, can be of definite interest for controlling light-emitting devices (including here injection lasers).[2] In fact, the ISR makes it possible to vary the size of the light-emitting region [and, if the gates are supplied asymmetrically, also its position in the interval $(-l, l)$]. By varying the current density and the carrier concentration in the conductive region in this case, we can (1) cross the threshold of the laser regime, and (2) vary the ratio between the radiative and nonradiative recombination channels. Changing the position and size of the light-emitting region has the possibility of varying the interaction of the charge carriers with the light field, varying the frequency of the radiation in the laser regime, modulating the output radiation, etc.

Therefore, a detailed study of the ISR (not only static properties, but also transition processes) can be of considerable practical interest.

2. Here we shall propose a theory of the potential distribution, current density, and carrier concentrations in the ISR for a $p^+-n-p-n^+$ structure that satisfies the following set of conditions:

a. The thicknesses of the internal regions w_p and w_n are small:

$$\omega_{n,p} \ll 1. \tag{4}$$

Such inequalities are not satisfied in actual GTO thyristors. In the controlling p base, the weaker condition $w_p < l$ is more often satisfied, but in the wide controlled n base (as a consequence of the requirement of high voltage), we have an inequality which is the opposite of condition (4). However, in actual light-emitting and laser thyristor structures, which are currently very widely used (see, for example, Refs. 3–10), condition (4) generally satisfied, with a good margin. Inequality (4) makes it possible to use a quasi-one-dimensional approach to describe the carrier distribution in the bases.

b. Each internal base is assumed to be so highly doped that a low-injection-level regime is realized in them at current densities of interest to us; i.e., the condition $n \ll p$ holds in the p base, and the opposite condition holds in the n base. To make our estimates easier, we also always assume both bases to be homogeneously doped. The smallness of the injection levels allows us to ignore the voltage drop across the bases in the x direction, assuming that the entire drop is concentrated at the three $p-n$ junctions (i.e., $\varphi = \varphi_1 + \varphi_2 + \varphi_3$). The potential of the bases thus depends on y: $\varphi_p(y) = \varphi_1(y)$, $\varphi_n(y) = \varphi_1(y) + \varphi_2(y) = \varphi(y) - \varphi_3(y)$.

c. The functions of the bases are separate. In general, we assume that one of the bases—the controlling p base—is more highly conductive (i.e., much more highly doped) and at the same time thinner (and therefore is characterized by a higher gain). The second base (in this case the n base), with lower conductivity and lower gain, makes the chief contribution to the recombination.

The choice of the n^+ cathode as the common electrode and of the p base as the controlling electrode is based on the dominant, routinely used structure of a silicon power thyristor. In the case of a controllable light-emitting $p^+-n-p-n^+$ diode made from GaAs or other III–V material (or alloy), the conductivity type of the layers can be something else. Since it is preferred to use a base made from a p-type material in the case of an ordinary LED, it is more suitable in the case of a $p^+-n-p-n^+$ diode to use the p base as the main light-emitting base. The n base must then assume the function of the controlling base, while the p^+ anode must become the common electrode. The transition to such a structure from the case considered below—the traditional structure—is obvious.

3. The basic system of equations, which determines the steady-state potential distribution of the two bases (φ_p and φ_n), consists of two continuity equations of the currents in these bases:

$$\frac{d\mathcal{J}_g^{(p)}}{dy} = j_1 - j_2, \tag{5}$$

$$\frac{d\mathcal{J}_g^{(n)}}{dy} = j_2 - j_3, \tag{6}$$

where $\mathcal{J}_g^{(p)}$ is the total current (per unit length in the z direction) flowing in the p base in the y direction, $\mathcal{J}_g^{(n)}$ is the same current in the n base, and $j_{1,2,3}$ are the densities of the currents flowing through each of the three $p-n$ junctions of the structure. Since we are considering only small injection levels

$$\mathcal{J}_g^{(p,n)} = -\sigma_{p,n} \frac{d\varphi_{p,n}}{dy}, \tag{7}$$

which where the $\sigma_{p,n}$ are the longitudinal conductivities of the bases, which are independent of the biases and which are determined by the majority carriers, the densities $j_{1,2,3}$ are determined by the voltages $\varphi_{1,2,3}(y)$, so that, to obtain the corresponding functionals, it is necessary to solve the problem of the two-dimensional distribution of minority carrier concentrations in the bases. In the p base, this equation has the form

$$\frac{\partial^2 n}{\partial x^2} + \frac{\partial^2 n}{\partial y^2} + \frac{\mu}{D}\frac{\partial}{\partial y}\left((n+n_0)\frac{d\varphi_p}{dy}\right) = \alpha_p^2 n, \tag{8}$$

where n is the nonequilibrium part of the electron concentration [i.e., the total concentration equals $n_0 + n(x,y)$, where n_0 is the equilibrium concentration]; $\alpha_p^2 = (D_n^{(p)}\tau^{(p)})^{-1}$ is the square inverse diffusion length; $D_n^{(p)}$ is the electron diffusion coefficient in the p base; $\tau^{(p)}$ is

the electron lifetime, which is independent of n because of the small injection levels for any recombination mechanisms; and μ/D is the ratio of the electron mobility in the p base to the diffusion coefficient, determined by the Einstein relation, $\mu/D=e/T$.

The equation analogous to Eq. (8) for the holes in the n base has the form

$$\frac{\partial^2 p}{\partial x^2}+\frac{\partial^2 p}{\partial y^2}-\frac{\mu}{D}\frac{\partial}{\partial y}\left((p+p_0)\frac{d\varphi_n}{dy}\right)=\alpha_n^2 p, \tag{9}$$

where the meaning of the new parameter α_n^2 is obvious.

Equations (8) and (9) include field terms associated with the $\varphi_{n,p}(y)$ dependence, but there are no analogous terms in the fields $E_x=-\partial\varphi/\partial x$. For small injection values, it would seem that it is possible to ignore the latter terms on the left-hand sides of Eqs. (8) and (9), retaining only the diffusion fluxes in them. However, the boundary conditions on the $p-n$ junctions,

$$n_1=n_0[\exp(e\varphi_p/T)-1], \tag{10}$$

$$n_2=n_0\{\exp[e(\varphi_p-\varphi_n)/T]-1\}, \tag{11}$$

$$p_2=p_0\{\exp[e(\varphi_p-\varphi_n)/T]-1\}, \tag{12}$$

$$p_3=p_0\{\exp[e(\varphi-\varphi_n)/T]-1\}, \tag{13}$$

($n_{1,2}$ and $p_{2,3}$ are the charge-carrier concentrations near $p-n$ junctions *1*, *2*, and *3*) show that the third terms on the left-hand sides of Eqs. (8) and (9) are of the same order of magnitude as the second terms, while disregarding these terms and the others corresponds to the quasi-one-dimensional approximation in the bases: the y coordinate appears only as a parameter in the boundary conditions. Solving Eqs. (8) and (9) and obtaining the $n(x,y)$ and $p(x,y)$ distributions makes it possible to compute the normal components of the diffusion fluxes at the $p-n$ junctions:

$$j_{n1}=-D_n^{(p)}\left.\frac{dn}{dx}\right|_{x=0},\quad j_{n2}=-D_n^{(p)}\left.\frac{dn}{dx}\right|_{x=\omega_p}, \tag{14}$$

$$j_{p2}=-D_p^{(n)}\left.\frac{dp}{dx}\right|_{x=\omega_p},\quad j_{p3}=-D_p^{(n)}\left.\frac{dp}{dx}\right|_{x=\omega_p+\omega_n}. \tag{15}$$

These fluxes are functionals of φ_p and φ_n, since the $n(x,y)$ and $p(x,y)$ distributions are obtained with the boundary conditions of Eqs. (10)–(13). For the current densities $j_{1,2,3}$ that appear in Eqs. (5) and (6) we have

$$j_1=-ej_{n1}+j_{1g}(\varphi_p), \tag{16}$$

$$j_2=ej_{p2}-ej_{n2}+j_{2g}(\varphi_p-\varphi_n), \tag{17}$$

$$j_3=ej_{p3}+j_{3g}(\varphi-\varphi_n). \tag{18}$$

The last terms on the right-hand sides of Eqs. (16)–(18) are the generation–recombination currents directly in the layers of the volume charge of the $p-n$ junctions, given in the form of direct functions of the voltages on these junctions.

The currents $j_{1,2,3}$ computed from Eqs. (16)–(18) and used in Eqs. (5) and (6) make it possible, using Eq. (7), to obtain a closed system of equations for determining $\varphi_{p,n}$. Equations (5) and (6) are solved with the boundary conditions at $y=\pm l$. These are either the specified-current conditions

$$\mathscr{J}_g^p(\pm l)=\pm\mathscr{J}_g^{(\pm)}, \tag{19}$$

where $\mathscr{J}_g^{(+)}=\mathscr{J}_g^{(-)}$ in the case of symmetric supply, or the specified-potential conditions

$$\varphi_p^{(\pm l)}=\varphi_g^{(\pm)}. \tag{20}$$

In each case we should add

$$\mathscr{J}_g^{(n)}(\pm l)=0 \tag{21}$$

to these conditions, since the n base (according to the structure model, see Fig. 1) has no current contacts.

If the quasi-one-dimensional approximation is used in the bases, the system of equations and boundary conditions introduced above is complete. If the treatment of the carrier distributions in the bases is to be two-dimensional, the system of equations should be supplemented with

$$j_{ny}=-D_n^{(p)}\frac{dn}{dy}+\mu_n^{(p)}(n+n_0)\left.\frac{d\varphi}{dy}\right|_{y=\pm l}=0$$

in the p base and an analogous condition in the n base. Since the inequalities are assumed to be $l\gg w_{p,n}$, the contribution of these supplementary conditions should not be of any significance in the regimes of interest to us.

Ordinarily it is not the voltage φ that is recorded on the structure but the total current $\mathscr{J}_a=\int_{-l}^{l}j_3(y)dy$, which makes it possible to compute φ for each given set of $\mathscr{J}_g^{(\pm)}$ or $\varphi_g^{(\pm)}$.

4. The possibility of using the quasi-one-dimensional approach is especially promising. Let us consider this possibility in more detail. From Eq. (8) without the second and third terms on the left-hand side, we have

$$n=n_1(y)\ \sinh\alpha_p(\omega_p-x)/\sinh\alpha_p\omega_p + n_2(y)\ \sinh\alpha_p x/\sinh\alpha_p\omega_p. \tag{22}$$

Likewise, from Eq. (9) we have

$$p=p_2(y)\ \sinh\alpha_n(\omega_n-x)/\sinh\alpha_n\omega_n + p_3(y)\ \sinh\alpha_n x/\sinh\alpha_n\omega_n; \tag{23}$$

in Eq. (23), we first transposed the x readings into the n region at the boundary with $p-n$ junction *2*. Equations (22) and (23) make it possible to implement Eqs. (14) and (15), to obtain currents $j_{1,2,3}$ from Eqs. (16)–(18), and to compute Eqs. (5) and (6) in explicit form:

$$\sigma_p\frac{d^2\varphi_p}{dy^2}=j_{2g}(\varphi_p-\varphi_n)-J_{1g}(\varphi_p) + eD_n^{(p)}\alpha_p\tanh\left(\frac{\alpha_p\omega_p}{2}\right)(n_1+n_2) + eD_p^{(n)}\alpha_n[p_2\cosh(\alpha_n\omega_n) - p_3]/\sinh\alpha_n\omega_n, \tag{24}$$

$$\sigma_n \frac{d^2\varphi_n}{dy^2} = j_{3g}(\varphi-\varphi_n) - j_{2g}(\varphi_p-\varphi_n)$$
$$- eD_p^{(n)}\alpha_n \tanh\left(\frac{\alpha_n\omega_n}{2}\right)(p_2+p_3)$$
$$- eD_n^{(p)}\alpha_p[n_2\cosh(\alpha_p\omega_p) - n_1]/\sinh\alpha_p\omega_p, \quad (25)$$

where $n_{1,2}$ and $p_{1,2}$ are explicit functions of φ_p and φ_n according to Eqs. (10)–(13). We thus obtain a system of two, coupled, nonlinear, second-order differential equations, which determine, along with the boundary conditions of Eqs. (21) and (19) [or Eq. (20)], the $\varphi_p(y)$ and $\varphi_n(y)$ distributions for a given $\mathscr{J}_a$ value.

In its complete form, the system of equations (24) and (25) is an object to be studied by computer. Assuming, however, that the main contribution both to current $\mathscr{J}_a$ and to current $\mathscr{J}_g$ comes from the region with high φ_p, $\varphi_p-\varphi_n$, and $\varphi-\varphi_n$ values, we can disregard the unities next to the exponentials in Eqs. (10)–(13) and also the currents $j_{1,2,3g}$ next to the diffusion currents on the right-hand sides of Eqs. (24) and (25). We can then rewrite Eqs. (24) and (25) in the form

$$\frac{d^2\psi_p}{dy^2} = \beta_p^2(e^{\psi_p}+e^{\psi_p-\psi_n}) + \beta_n^2[(1+\gamma_n)e^{\psi_p-\psi_n} - \gamma_n e^{\psi-\psi_n}], \quad (26)$$

$$\xi\frac{d^2\psi_n}{dy^2} = -\beta_n^2(e^{\psi_p-\psi_n}) - \beta_p^2[(1+\gamma_p)e^{\psi_p-\psi_n} - \gamma_p e^{\psi_p}], \quad (27)$$

where

$$\psi_{p,n} = e\psi_{p,n}/T, \quad \xi = \sigma_n/\sigma_p,$$
$$\beta_p^2 = \frac{e^2D_n^{(p)}\alpha_p n_0}{\sigma_p T}\tanh\left(\frac{\alpha_p\omega_p}{2}\right),$$
$$\beta_n^2 = \frac{e^2D_p^{(n)}\alpha_n p_0}{\sigma_p T}\tanh\left(\frac{\alpha_n\omega_n}{2}\right),$$
$$\gamma_n = 1/2\sinh^2\left(\frac{\alpha\omega_n}{2}\right), \quad \gamma_p = 1/2\sinh^2\left(\frac{\alpha_p\omega_p}{2}\right).$$

5. The controlling base actually possesses a substantially higher longitudinal conductivity than the controlled base; therefore, the parameter ξ is small, $\xi \ll 1$. This makes it possible to obtain a partial solution of the system of equations (26) and (27) which does not explicitly use the condition given in Eq. (21) (since we ignore the current through the n base everywhere), completely omitting the left-hand side of Eq. (27); in this case we have

$$e^{-\psi_n} = \gamma_p e^{\psi_p}\left[(1+\gamma_p)e^{\psi_p} + \frac{1}{\varkappa^2}(e^{\psi_p}+e^{\psi})\right]^{-1} \quad (28)$$

and

$$\frac{d^2\psi_p}{dy^2} = \beta_p^2 e^{\psi_p}\left(1+\gamma_p\frac{(1+\gamma_n)e^{\psi_p}-\gamma_n e^{\psi}+\varkappa^2 e^{\psi_p}}{e^{\psi_p}+e^{\psi}+\varkappa^2(1+\gamma_p)e^{\psi_p}}\right), \quad (29)$$

where $\varkappa = \beta_p/\beta_n$. Equation (29) is integrated in quadratures, which makes it possible to study the given limiting case analytically.

In the case of a uniform distribution of the potential φ_p and the current density, setting the right-hand side of Eq. (29) to zero, we have

$$e^{\psi_p} = e^{\psi_p^{(0)}} = e^{\psi}\frac{\gamma_p\gamma_n - 1}{\gamma_p\gamma_n+\gamma_p+1+\varkappa^2(1+2\gamma_p)}. \quad (30)$$

In other words, the solution that we need must satisfy the condition $\gamma_p\gamma_n > 1$ or

$$\cosh^{-1}(\alpha_p\omega_p) + \cosh^{-1}(\alpha_n\omega_n) > 1. \quad (31)$$

This is the well-known condition for the open state of a thyristor. Using Eq. (30), it is convenient to write Eq. (29) in dimensionless form,

$$\frac{d^2\chi}{d\eta^2} = \frac{e^{\chi}(e^{\chi}-1)}{1+e^{\chi}/A}, \quad (32)$$

where $\eta = \beta_p y\sqrt{\gamma_p\gamma_n - 1}\exp(\psi_p^{(0)}/2)$, $\chi = \psi_p - \psi_p^{(0)} < 0$, and $A = (1+\gamma_p+\gamma_p\gamma_n+\varkappa^2(1+2\gamma_p))/(\gamma_p\gamma_n-1)\times(1+\varkappa^2(1+\gamma_p))$. When $\varkappa^2 \ll 1$, we have $A \simeq (1+\gamma_p+\gamma_p\gamma_n)/(\gamma_p\gamma_n-1)$. We should point out that Eq. (32) is much simpler when $A \gg 1$. This condition is satisfied, first, in the case of small supercriticality, i.e., when $\gamma_p\gamma_n - 1 \ll 1$; second, for moderate supercriticality ($\gamma_p\gamma_n - 1 \simeq 1$), which is attained when $\gamma_p \gg 1$ (and, consequently, $\gamma_n \ll 1$), i.e., when there is a sharp difference in the gains of the bases.

Equation (32) has two homogeneous solutions, for which $d^2\chi/d\eta^2 \simeq 0$:

(1) $\chi = 0$, i.e., $\psi_p = \psi_p^{(0)}$; this is a homogeneous conductive region with no transverse electric field ($d\chi/d\eta = 0$), which represents a thyristor in the open state.

(2) $\chi \to -\infty$ (i.e., $\psi_p \to -\infty$); this is a blocked region with a homogeneous transverse field,

$$\frac{d\chi}{d\eta} = -\frac{\mathscr{J}_g}{I_0\beta}e^{-\psi_p^{(0)}/2}, \quad (33)$$

where $I_0 = \sigma_p T/e$, and $\beta = \beta_p\sqrt{\gamma_p\gamma_n - 1}$. In what follows, we shall use $\mathscr{J}_g$ to mean the absolute value of the blocking current.

Let us consider such an nonuniform distribution $\chi(\eta)$, which in the limit $\eta \to -\infty$, undergoes a transition to the homogeneous state $\psi_p = \psi_p^{(0)}$ (with no transverse field or transverse current), and in the limit $\eta \to \infty$, it undergoes a transition to a distribution with the constant field, given by Eq. (33), and a constant current through the p base. We should point out that this current is "generated" in the intermediate region with a nonuniform field distribution, $d\chi/d\eta$, which we are considering. Integrating Eq. (32) once, we obtain

$$\frac{1}{2A}\left(\frac{d\chi}{d\eta}\right)^2 = (1+A)\ln\left(\frac{1+A}{e^{\chi}+A}\right) - 1 + e^{\chi}. \quad (34)$$

When $\chi \to -\infty$, it follows from this equation that

$$\pm\frac{d\chi}{d\eta}=(2A(1+A)\ \ln(1+1/A)-2A)^{1/2}\equiv C. \tag{35}$$

In other words, using Eq. (33), we obtain

$$\mathscr{J}_g=C\beta I_0\ e^{\psi_p^{(0)}/2}. \tag{36}$$

When $A\gg 1$, it follows from Eq. (34) that

$$\frac{d\chi}{d\eta}=\pm(1-e^{\chi}), \tag{34'}$$

and for C we have $C=1$.

Since for the current density in a homogeneous conductive region (where $j_a=j_1=j_2=j_3$) we have

$$j_a^{(0)}=\beta_p^2 A_1 I_0\ e^{\psi_p^{(0)}}, \tag{37}$$

where

$$A_1=\frac{(1+\gamma_p)(1+2\gamma_n)+\varkappa^2(1+\gamma_n)(1+2\gamma_p)}{1+2\gamma_n+\varkappa^2(1+\gamma_n+\gamma_n\gamma_p)},$$

using Eq. (36), we obtain

$$\mathscr{J}_g=C_1(I_0 j_a^{(0)})^{1/2}, \tag{38}$$

where $C_1=\{(2A/A_1)(\gamma_p\gamma_n-1)[(1+A)\ln(1+1/A)-1]\}^{1/2}$. When $\varkappa^2\ll 1$, we have $A_1\simeq 1+\gamma_p$; if it is also assumed that $A\gg 1$, then $C_1=[(\gamma_p\gamma_n-1)/(1+\gamma_p)]^{1/2}$. The specific form of the transition layer can be obtained by integrating Eq. (34), but the spatial scale which describes its thickness can be obtained, for example, by linearizing Eq. (32) for small χ values. In this case, it is given by the length

$$y_c=\frac{e^{-\psi_p^{(0)/2}}(1+1/A)^{1/2}}{\beta_p(\gamma_p\gamma_n-1)^{1/2}}=C_2(I_0/j_a^{(0)})^{1/2}, \tag{39}$$

where $C_2=[A_1(1+1/A)/(\gamma_p\gamma_n-1)]^{1/2}$. When $A\gg 1$, it follows from Eq. (34') that

$$e^{\eta-\eta_0}=e^{-\chi}-1, \tag{40}$$

where η_0 is a certain arbitrary quantity that characterizes the position of the transition layer in the interval $(-\infty, +\infty)$. When $|\chi|>1$, a linear dropoff is obtained for χ; i.e., $\chi=\eta_0-\eta$. When the conditions $A\gg 1$ and $\varkappa^2\ll 1$ are satisfied at the same time, we have

$$j_c=j_1\simeq\beta_p^2 I_0\ e^{\psi_{(0)}^{p}}(\gamma_p+1)e^{\chi}, \tag{41}$$

$$j_a=j_3=j_2\simeq\beta_p^2 I_0 e^{\psi_p^{(0)}}(\gamma_p+1)e^{\chi}\left[1+\frac{\gamma_p\gamma_n-1}{\gamma_p+1}(1-e^{\chi})\right]. \tag{42}$$

By combining Eqs. (41) and (42) with Eq. (40) we have obtained simple dependences that determine the spatial distribution of the anode current density $j_a(\eta)$, the cathode current density $j_c(\eta)$, and the "generation" rate of the gate current $j_a(\eta)-j_c(\eta)$ in the transition layer.

It is natural that the solution described above in the form of two regions with a transition layer between them can be fitted into the half-strip with a width l when a number of strong inequalities are satisfied. Most importantly, the condition $y_c\ll l$ should be satisfied, or, using Eq. (38), the condition

$$C_1C_2I_0/\mathscr{J}_g\ll l \tag{43}$$

should be satisfied; the product C_1C_2 virtually always differs only slightly from 1. As $\mathscr{J}_g$ increases, and, for a rigorously given $\mathscr{J}_a$ value, as a consequence of the increase of current density j_a in the conductive region, its effective size is compressed,

$$x_c=\mathscr{J}_a/j_a^{(0)}. \tag{44}$$

In this case, the characteristic size y_c, according to Eq. (39), decreases more slowly than x_c, and the condition

$$y_c\ll x_c, \tag{45}$$

beginning with a certain value of $\mathscr{J}_g$, breaks down; to satisfy inequality (45), it is necessary to satisfy the inequality

$$\mathscr{J}_g\ll\mathscr{J}_a C_1/C_2. \tag{46}$$

In other words, an inhomogeneous distribution, which includes an almost homogeneous conductive region and a pronounced depletion region, is formed in the range of $\mathscr{J}_g$ values determined by conditions (43) and (46):

$$C_1C_2I_0/l\ll\mathscr{J}_g\ll\mathscr{J}_a C_1/C_2. \tag{47}$$

This range is itself formed for a sufficiently large $\mathscr{J}_a$ value

$$\mathscr{J}_a\gg C_2^2 I_0/l. \tag{47'}$$

Let us estimate, first, the possibility of satisfying inequality (47') and then the existence of the necessary interval of $\mathscr{J}_g$ values. We thus will estimate the value of $\mathscr{J}_0/l$. For a conductivity of the controlling base of $\sim 10^2\ (\Omega\cdot\text{cm})^{-1}$ and a base thickness w_p of $\sim 10^{-4}$ cm, we have $\sigma_p\simeq 10^{-2}\ \Omega^{-1}$. When $T\simeq 300$ K, we obtain $I_0=2.5\times 10^{-4}$ A. When $l=50\mu$m, we have $I_0/l=5\times 10^{-2}$ A/cm. Setting the constant $C_2^2\simeq 1$ (which is reasonable when the supercriticality $\gamma_p\gamma_p-1$ is not too small), we obtain condition (44) in the form $j_a^{(0)}=\mathscr{J}_a/2l\gg I_0/2l^2=5$ A/cm^2. Since it is required to satisfy the condition of small injection levels, there is a certain upper limit of possible j_a values. When the acceptor concentration in the p base exceeds 10^{18} cm^{-3} and its thickness is $\sim 10^{-4}$ cm, this limiting density exceeds 10^3 A/cm^3, so that we have some range in which the calculation being developed is applicable.

6. Let us consider the character of the $j_a(y)$ distribution when $\mathscr{J}_g$ is outside the interval of inequalities (47). For small $\mathscr{J}_g$ values corresponding to an inequality opposite to inequality (43), we encounter the situation symbolized by inequality (1). In fact, the product $C_1C_2=\{2(1+A)\times[(1+A)\ln(1+1/A)-1]\}^{1/2}$ is rigorously equal to unity when $A\to\infty$, but, as pointed out above, differs little from unity virtually everywhere (even for $A=1$, which is the lower limit of this quantity). When inequality (1) is satisfied, as was pointed out, we can ignore the dependence of all the concentrations and potentials on y by considering the homogeneous state. In this state we have

$$\mathscr{J}_g = lA\beta^2 I_0\, e^{\psi_p}\frac{1-e^{\chi}}{A+e^{\chi}}, \tag{48}$$

$$\mathscr{J}_a = 2l\beta_p^2 I_0\,\frac{\gamma_p\gamma_n}{1+\varkappa^2(1+\gamma_p)}\,e^{\psi_p}\frac{B-e^{\chi}}{A+e^{\chi}}, \tag{49}$$

where $B=A(1+1/\gamma_n)[1+\varkappa^2(1+\gamma_p)]>1$. It is possible to use Eqs. (48) and (49) to express $\exp(\psi_p)$ and $\exp(\chi)$ in terms of $\mathscr{J}_g$ and $\mathscr{J}_a$. As a result, we have

$$e^{\chi}=\frac{B_1-B(2\mathscr{J}_g/\mathscr{J}_a)}{B_1-(2\mathscr{J}_g/\mathscr{J}_a)}, \tag{50}$$

where $B_1=1+[1+j_p+\varkappa^2(1+2\gamma_p)]/j_nj_p>1$. Increasing $\mathscr{J}_g$ for a given $\mathscr{J}_a$ causes a decrease of $\exp(\chi)$, which reaches the zero mark when

$$\frac{2\mathscr{J}_g}{\mathscr{J}_a}=\frac{B_1}{B}=\frac{(\gamma_p\gamma_n-1)}{\gamma_p(1+\gamma_n)}. \tag{51}$$

This limiting ratio $2\mathscr{J}_g/\mathscr{J}_a$ corresponds to the limiting value

$$e^{\psi_p}=\frac{\mathscr{J}_a}{2lI_0\beta_p^2\gamma_p(1+\gamma_n)}. \tag{52}$$

The attainment of this limit is possible only for fairly small $\mathscr{J}_a$, for which, instead of Eq. (47′), the inequality opposite in sense is satisfied.

When condition (47′) is satisfied and $\mathscr{J}_g$ reaches the values given by the right-hand side of inequality (47), the conductive region is compressed to the order of magnitude of its wall thickness and loses homogeneity. It is obvious in this case that current $\mathscr{J}_g$ already controls the current density j_a not only by varying the size, but also "directly"—by selecting the part of the anode current in the base electrode. In this case $\psi_p(0)$ is less than the $\psi_p^{(0)}$ value determined by Eq. (30) from the given value of the total voltage ψ on the structure. Assuming that $\exp[\psi_p(0)]\ll\exp(\psi_p^{(0)})$, we easily obtain the following expressions after integrating Eq. (32) and computing the total current $\mathscr{J}_a$:

$$\mathscr{J}_g=\beta I_0(2\,e^{\psi_p(0)})^{1/2},$$

$$\mathscr{J}_a=1\beta I_0\,\frac{\gamma_p(1+\gamma_n)}{\gamma_p\gamma_n-1}\,(2e^{\psi_p(0)})^{1/2};$$

i.e., some limiting relationship arises between $\mathscr{J}_g$ and $\mathscr{J}_a$[see inequality (3)]:

$$I_g=\mathscr{J}_M(\mathscr{J}_a)=\frac{\mathscr{J}_a}{2}\,\frac{\gamma_p\gamma_n-1}{\gamma_p(1+\gamma_n)}. \tag{53}$$

It is easy to see that the right-hand side of this equation is not much different from the right-hand side of the strong inequality (46). There are no stationary states with a given total current $\mathscr{J}_a$ for $\mathscr{J}_g$ values whicht exceed the right-hand side of Eq. (53). We should point out that Eq. (53) does not differ from Eq. (51), which was obtained for small $\mathscr{J}_g$ and $\mathscr{J}_a$ values.

7. The above calculations and estimates make it possible to describe a fairly unified picture of how the blocking gate current $\mathscr{J}_g$ affects the distribution of the anode current density j_a in a four-layer structure with a given total current $\mathscr{J}_a$. While this total current is small and the inequality which is opposite in sense to inequality (47′) is satisfied, the action of $\mathscr{J}_a$ nowhere causes a significant change in the current density j_a, which roughly retains its homogeneity. The action of $\mathscr{J}_g$ causes only an increase of the total voltage ψ on the region. The increase in $\mathscr{J}_g$ is restricted by the limiting relationship given by Eq. (51): steady-state values of $\mathscr{J}_g$ which exceed this level are impossible for the given value of $\mathscr{J}_a$.

When $\mathscr{J}_a$ increases and inequality (47′) is satisfied, the action of $\mathscr{J}_g$ changes sharply. This current accounts for an essentially heterogeneous picture of the $j_a(y)$ distribution; the conductive region is compressed—its half-dimension decreases as $\mathscr{J}_g$ increases:

$$x_c=C_1^2\,\frac{I_0\mathscr{J}_a}{\mathscr{J}_g^2}, \tag{54}$$

while the current density in it increases accordingly [see Eq. (38)]. The total voltage ψ also increases accordingly. We are mainly interested in precisely this heterogeneous picture, in which the size of the conducting region and its current density can be controlled. Even in this case, the possible increase of $\mathscr{J}_g$ is not infinite; in the stationary case, this current can attain only the maximum value of Eq. (53), which corresponds to the limiting compression of the conducting region. When it approximates this value, the criteria for the approximation of small injection levels used in the above calculations and estimates break down. This breakdown is apparently substantial only for the controlling p base, since, when $\xi\ll 1$ and $\varkappa\ll 1$, the controlled n base appears in our formulas only in the form of the argument $(\alpha_n w_n)$ in the hyperbolic functions (i.e., ultimately as the gain of this base with respect to current). This, of course, relates to the rather thin n bases in the regime in which the main voltage drop occurs at the $p-n$ junctions. The criteria used above for the quasi-one-dimensional approach can also break down for the limiting compression of the conductive region.

Let us illustrate what has been said by means of graphs. Figures 2 and 3 show the spatial distribution of the anode current density λ_a for two different structures as a function of the gate current. We use the dimensionless spatial variable $\zeta=y/l$, which allows us in all cases to use the interval (0,1).

The graphs were drawn by solving Eq. (32) for the given $\mathscr{J}_g$ and $\mathscr{J}_a$ values. Each $\lambda_a=\lambda_a(\zeta)$ curve is determined by the five parameters already introduced, γ_n, γ_p, $\varkappa^2$, and also $\Lambda_a=\mathscr{J}_a l/2\mathscr{J}_0$ and $\Lambda_g=\mathscr{J}_g l/2\mathscr{J}_0$. The anode current density is normalized in such a way that $\int_0^1\lambda_a(\zeta)d\zeta = 1$, i.e., $\lambda_a=lj_a/\mathscr{J}_a$. In Fig, 2, a family of $\lambda_a(\zeta))$ dependences for a given Λ_a value and various Λ_g values is constructed when condition (47′) is definitely satisfied and a sufficient range of $\mathscr{J}_g$ values, which satisfy conditions (47), exists. Both the blocked regions and the homogeneous conductive regions, which are separated by comparatively narrow transition layers clearly seen in this range. In contrast with Fig. 2, Fig. 3 illustrates the situation in which, condition (47′) is satisfied, though not too clearly. Even in this case, as $\mathscr{J}_g$ increases, we obtain essentially heterogeneous states, but the range with homogeneous narrowed conductive regions is missing.

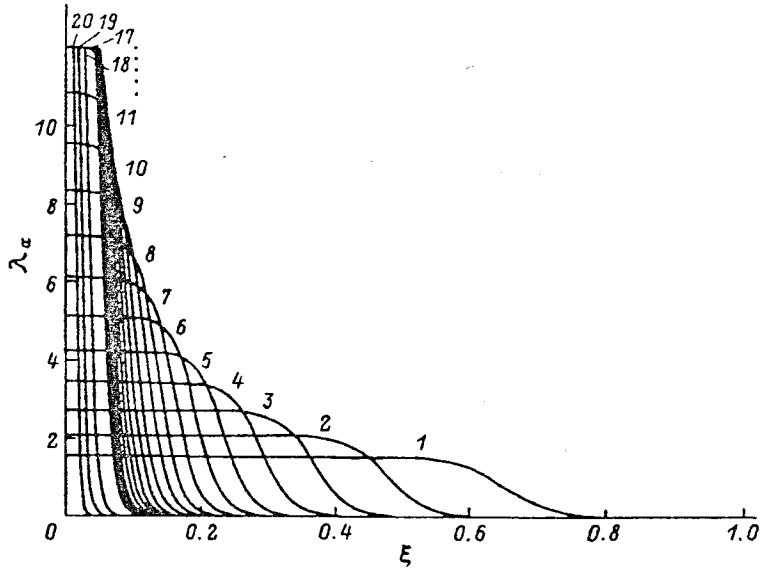

FIG. 2. Anode current density distribution $\lambda_a(\zeta)$ for various gate currents Λ_g. The curves with numbers $N=1-17$ are constructed for currents $\Lambda_g(N)=30+5(N-1)$, and the curves with $N=18$, 19, and 20 are constructed for $\Lambda_g=150$, 200, and 275, respectively. (The numbers on curves 12–16 in the figure are not shown.) The value $\Lambda_g=275$ corresponds to $\mathscr{J}_g=\mathscr{J}_M(\mathscr{J}_a)$. The other parameters are $\Lambda_a=500$, $\gamma_p=3$, $\gamma_n=2$, $\varkappa^2=0.5$, and $\zeta=0$.

We should point out that the values of the parameters γ_n, γ_p, and $\varkappa^2$ used for the curves in Figs. 2 and 3 do not involve the simplifying approximations $A\gg 1$ or even $\varkappa^2\ll 1$. In Fig. 4, the dependences $\lambda_a(\zeta)$ and $\lambda_c(\zeta)=lj_c/\mathscr{J}_a$ are shown simultaneously for several values of $\mathscr{J}_a$ and $\mathscr{J}_g$. In addition to the case of an intermediate value of $\mathscr{J}_g$ with a homogeneous conductive region [Fig. 4a], we show the case of a large $\mathscr{J}_g$ with a compressed, inhomogeneous conductive region [Fig. 4b]. In the first case, the current $\mathscr{J}_g$ is "generated" only in the transition boundary layer. In the second case the entire conductive region is biased by the gate current. Figure 5 shows the potential distributions $\psi_p(\zeta)$ and $\psi_n(\zeta)$ for the same "sample" and current $\mathscr{J}_a$ as in Fig. 2; it also shows the levels of the total potential ψ, which also depends on $\mathscr{J}_g$. As $\mathscr{J}_g$ increases, the potential ψ increases, along with the potential $\psi_p(\zeta)$ in the conductive region, whereas the potential $\psi_n(\zeta)$ is retained as an invariant. It begins to vary when the conductive region loses its homogeneity.

Note the linear growth of the negative potential $\psi_p(\zeta)$ as the gate contact is approached. It is associated with the ohmic voltage drop as a result of the presence of current $\mathscr{J}_g$. The analogous linear increase of the positive potential $\psi_n(\zeta)$ has no physical meaning and is associated with a defect of the approximation (see Sec. 8). We should point out that this defect shows up only in the blocked region, where the current density $j_a(\zeta)$ is negligible.

8. The preceding treatment in no way allows for the finite value of the conductivity σ_n of the controlled n base. Taken together with the fact that the thermal carrier generation is ignored in Eqs. (26) and (27) [and, consequently, in Eq. (32)], this means that the n base reproduces with reversed sign the variation of the potential of the p base [see Eq. (28)]. As a result, a nonphysical growth of its floating potential arises in the blocked region of the controlled base, which blocks not only the middle (interbase) $n-p$ junction, but also the anode junction. This effect is clearly observed in Fig. 5, which shows the potential distribution of both bases. This unphysical behavior of potential ψ_n occurs in the region of small anode current density and therefore has a weak effect on the total results.

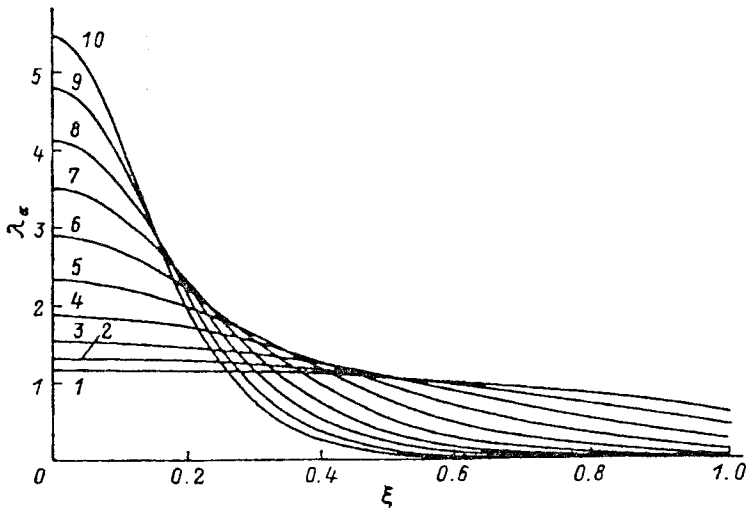

FIG. 3. Anode current density distribution $\lambda_a(\zeta)$ for various gate currents Λ_g. The curves with numbers $N=1-10$ correspond to currents $\Lambda_g(N)=2+(N-1)$. The value $\Lambda_g=11$ corresponds to $\mathscr{J}_g=\mathscr{J}_M(\mathscr{J}_a)$. The other parameters are $\Lambda_a=100$, $\gamma_p=3$, $\gamma_n=0.5$, $\varkappa^2=0.5$, and $\zeta=0$.

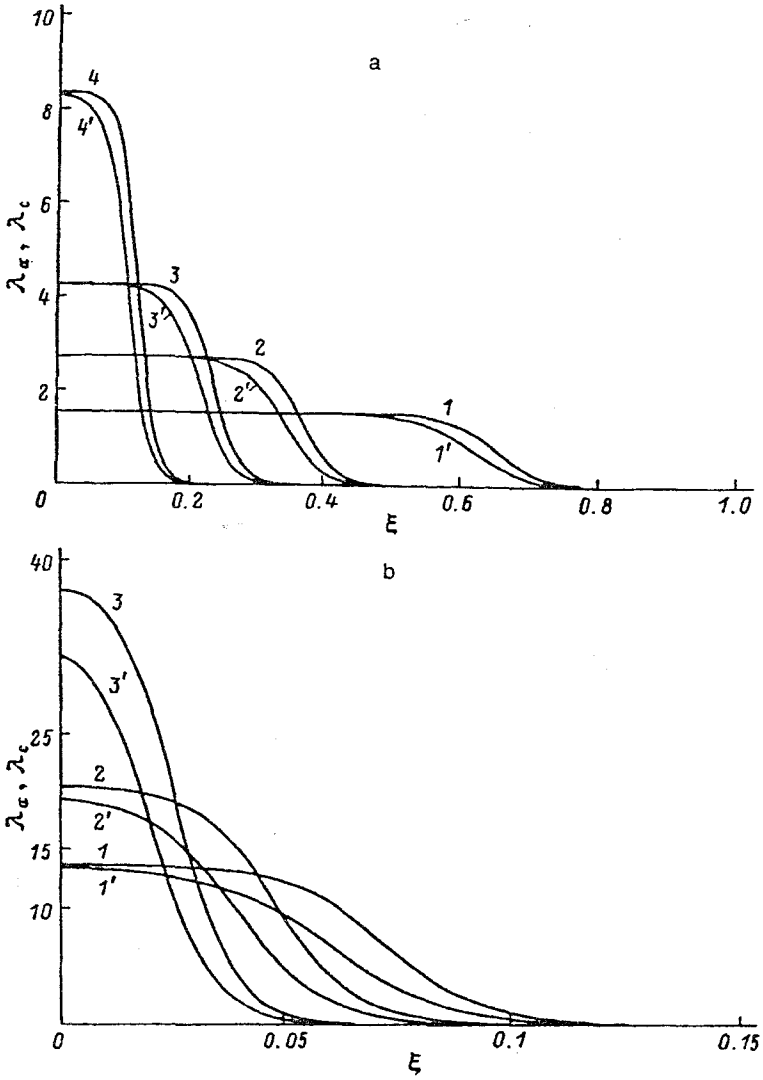

FIG. 4. Distributions of $\lambda_a(\zeta)$ (*1–4*) and $\lambda_c(\zeta)$ (*1′–4′*) for the same values of the sample parameters as in Fig. 2. *a*—$\Lambda_g(N)$ values: *1*—30; *2*—40; *3*—50; *4*—70. *b*—Λ_g values: *1*—90, *2*—110, *3*—150.

Here we shall attempt to roughly take into account the contribution of the low conductivity along the n base in the distribution of potential ψ_n and current density $j_a(y)$ in the blocked region. The presence of this conductivity results in a negative bias of the base of the p^+–n–p^- transistor, in which the forward- biased anode p^+–n junction is a hole emitter, while the reverse-biased middle n–p junction is a hole collector.

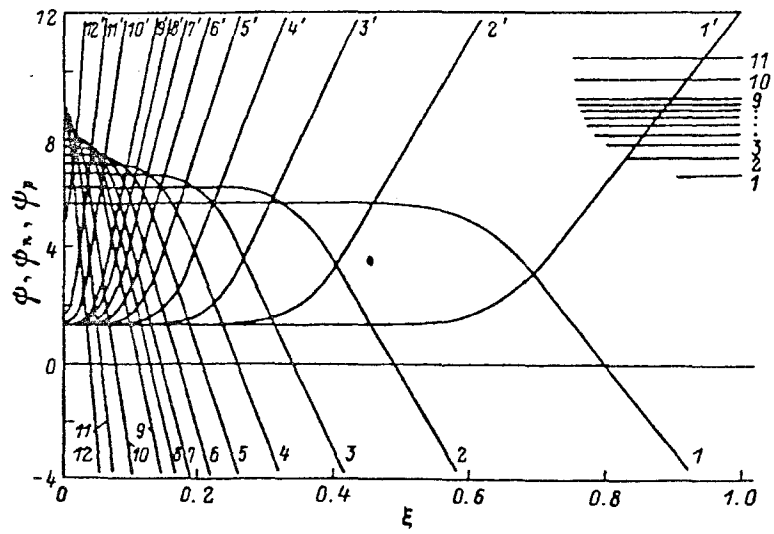

FIG. 5. Distributions of potentials $\psi(\zeta)$, $\psi(\zeta)$ (*1–12*) and $\psi_n(\zeta)$ (*1′–12′*) for the same parameters as in Fig. 2. The horizontal lines (*1–11*) at the top right show the corresponding Φ values. The curves numbered $N = 1-9$ correspond to the values $\Lambda_g(N) = 30 + 10(N-1)$, and those with $N = 10-12$ correspond to $\lambda_g = 150$, 200, and 275, respectively.

We restrict ourselves for simplicity to the more transparent case of $A \gg 1$. Here the right-hand side of Eq. (27) can be greatly simplified by dropping the terms with $\exp(\psi_p - \psi_n)$ in it. As a result, the equation acquires the form

$$\zeta' \frac{d^2\chi}{d\eta^2} = e^{\chi'} - e^{\chi}, \tag{55}$$

where $\zeta' = \zeta(\gamma_p\gamma_n - 1)/\gamma_p$, and $\chi' = \psi - \psi_n - \psi_p^{(0)} - \ln(\varkappa^2\gamma_p)$. We assume that the finiteness of quantity ζ has little effect on the dependence of $\chi(\eta)$, but strongly influences the $\chi'(\eta)$ dependences. This allows us to assume the function $\chi(\eta)$ in Eq. (55) as given [see Eq. (40)]:

$$\chi(\eta) \simeq -\ln[\exp(\eta - \eta_0) + 1]. \tag{56}$$

Assuming, as before, that $\chi'(\eta) \simeq \chi(\eta)$, we have

$$\frac{d\chi'}{d\eta} = -(1 + e^{-\eta + \eta_0})^{-1}; \tag{57}$$

i.e., the transition region "generates" not only the gate current of the controlling base, but simultaneously also a current in the controlled base, which is a factor of ζ lower than $\mathscr{J}_g$, which is directed in the opposite direction. In contract with the current $\mathscr{J}_g$, for which a special electrode exists—the gate—the current through the controlled base can enter only through the anode, which should be open even in the blocked region and through which should flow the forward current, part of which constitutes the mentioned current through the controlled base, while the other part enters the reverse-biased middle p–n junction. This means that we have in the blocked region an effective p^+–n–p transistor with the anode p^+–n junction as an emitter and the middle p–n junction as a collector. This transistor is biased with respect to the controlled n base by the base current, which flows through the conductive region.

To calculate the anode current distribution in such a transistor, we must solve Eq. (55) without the last term on the right-hand side (since it is negligible in the blocked region). The solution that satisfies the condition $d\chi'/d\eta|_{\eta=\eta_l} = 0$, where $\eta_l = \beta l \exp(\psi_p^{(0)}/2)$, is

$$e^{\chi'} = \frac{2\zeta'\Phi^2}{\eta_1^2 \cos^2[\Phi(\eta_l - \eta)/\eta_1]}, \tag{58}$$

where Φ is a constant of integration ($0 < \Phi < \pi/2$), and η_1 is the length of the blocked region (and, consequently, of the effective transistor), so that η varies in the interval $(\eta_l - \eta_1, \eta_l)$. The length of the blocked region is undoubtedly a conditional concept, whose precise quantitative meaning can be explained only after solving Eq. (55) in total form (which is left here for the future). Here we restrict the solution to equating the derivative $d\chi'/d\eta$, found from Eq. (58), to its value -1, which follows from Eq. (57) for $\eta = \eta_l - \eta_1$. For Φ we then obtain

$$2\Phi \tan \Phi = \eta_1. \tag{59}$$

For large blocked regions ($\eta_1 \gg 1$), $\Phi \simeq \pi/2$ follows from Eq. (59). Using this value in Eq. (58), we obtain for the edge point $\eta = \eta_l$ the expression

$$\exp[\chi'(\eta_l)] = \zeta'\pi^2/2\eta_1^2. \tag{60}$$

It is always possible to single out such large $\psi_p^{(0)}$ values (which increase with increasing current $\mathscr{J}_a$) that the value $\psi - \psi_n(\eta_l)$ is positive and the resulting estimates are legitimate.

9. Let us summarize the results.

a) We have demonstrated the possibility of steady-state compression of the conducting region in the structure under consideration. In the case of a large $\mathscr{J}_a l$ [see condition (47′)] this compression in a wide range of values of the gate current results from narrowing of the homogeneous conducting region, which is pressed toward the center of the structure by the peripheral blocked regions. In the case of moderate $\mathscr{J}_a l$ values, the gate current also substantially varies the anode current density distribution over the area of the anode, forming a heterogeneous structure with an inhomogeneous conductive region, whose size is on the order of its wall thickness.

b) In the case of large $\mathscr{J}_a l$, the size x_c of the homogeneous conductive region decreases with increasing $\mathscr{J}_g$ as $\mathscr{J}_g^{-2}$ [see Eq. (54)]. This result is in complete agreement with the analogous result of Ref. 1, which was obtained for a different thyristor structure under the same condition of low injection levels in the p base. Under the condition $\varkappa^2 \ll 1$, $\zeta \ll 1$ and when γ_n is independent of the current density, the controlled n base apparently does not have too critical an effect on the processes studied here. This makes it possible to hope that the gate can control the structure considered here, and that the assumptions made here concerning the n base can be substantially broadened.

It is easy to show that, when one goes from small to large injection levels in the controlling base (as in the case considered in Ref. 1), despite the variation of certain quantitative regularities, the qualitative character of the compression picture will remain the same, even though the theoretical treatment must be based on slightly different principles.

We wish to thank the Fund for Fundamental Research of the State Committee for Science and Technology of Ukraine for partial support of this study.

*Wayne State University, Detroit, Michigan 48202, USA

[1] Z. Gribnikov and A. Rothwarf, Solid State Electron. **37**, 135 (1994).
[2] Z. Gribnikov, V. Mitin, and A. Rothwarf, Proc. Conf. 1993 ISDRS **2**, 603 (1993).
[3] P. R. Claisee, G. W. Taylor, D. P. Doctor, and P. W. Cooke, IEEE Trans. Electron. Dev. **30**, 2523 (1992).
[4] P. W. Cooke, G. W. Taylor, and P. R. Claisee, IEEE Photon. Technol. Lett. **2**, 537 (1990).
[5] G. W. Taylor and P. W. Cooke, Appl. Phys. Lett. **56**, 1308 (1990).
[6] P. W. Cooke, G. W. Taylor, P. R. Claisee, and T. Y. Chang, J. Vac. Sci. Technol. B **8**, 367 (1990).
[7] D. L. Crawford, G. W. Taylor, and J. G. Simmons, Appl. Phys. Lett. **52**, 863 (1988).
[8] K. Kasahara, Y. Tashro, N. Hamao, M. Sugimoto, and T. Yanase, Appl. Phys. Lett. **52**, 679 (1988).
[9] Y. Tashiro, N. Hamao, M. Sugimoto, N. Takado, S. Asada, and T. Yanase, Appl. Phys. Lett. **54**, 329 (1989).
[10] K. Kurihara, T. Numai, I. Ogura, H. Kosaka, M. Sugimoto, and K. Kasahara, Japan. J. Appl. Phys. B **32**, 604 (1993).

Translated by W. J. Manthey