Electron bistability of doped semiconductors with polar optical scattering: reversible switching in CdTe:Cl at room temperature

N. V. Agrinskaya and V. I. Kozub

A. F. Ioffe Physicotechnical Institute, Academy of Sciences of the USSR, Leningrad
(Submitted 16 August 1990)
Zh. Eksp. Teor. Fiz. **99**, 929–944 (March 1991)

It is shown that bistable behavior of the electron system in doped compensated (by deep centers) semiconductors with polar optical scattering is possible in strong electric fields at room temperature. This bistability is manifested by an S-shaped current-voltage characteristic. The bistability mechanism is due to a combination of the runaway effect (typical of polar optical scattering), the impact ionization of deep centers, and the electron–electron scattering. A systematic analysis of the last effect is made going beyond the Landau collision integral approximation and allowing for the processes characterized by a large transferred energy. Such processes transfer electrons from the low- to the high-energy region and can ensure retention of a low-resistance state with a high characteristic electron energy and a correspondingly high degree of ionization of the deep level. A study is made of the influence of the side valleys that limit the runaway effect. The proposed model accounts for the experimental data on electrical breakdown and reversible switching in CdTe.

INTRODUCTION

The switching phenomena in strong electric fields (S-shaped current-voltage characteristics) have always attracted special interest because of the interesting physics and also due to technical applications. One can mention here the effects that occur in glassy chalcogenide semiconductors and are of thermal origin[1] or the various phenomena in diode structures.[2] The S-shaped current-voltage characteristics have been observed also in the case of strongly nonequilibrium low-temperature states of electron systems in bulk crystals of several materials (InSb, Ge, etc.), described, for example, in Ref. 3. Our aim will be to analyze nonlinear electrical properties of CdTe crystals which exhibit both impurity breakdown[4] and reversible switching at room temperature.[5,6]

The behavior of CdTe in strong electric fields has been investigated intensively on many occasions, particularly because the characteristic features of the energy band structure of this material facilitate the Gunn effect. However, although the Gunn effect is indeed observed in pure cadmium telluride,[7] in samples with high impurity concentrations the Gunn effect is replaced or supplemented by electrical breakdown[4] accompanied by an S-shaped current-voltage characteristic.[5] It was suggested in Refs. 4 and 5 that the Gunn effect is suppressed by the impurity scattering and by the interband breakdown. However, as pointed out in Ref. 5, this still does not account for the origin of the S-shaped current-voltage characteristic (maintaining a low-resistance state in weak fields).

It was reported in Ref. 6 that a similar reversible switching effect occurred in CdTe:Cl at room temperature. The concentration N_{Cl} of shallow donors with an energy level at $E_c - 15$ meV was $\sim 10^{17}$–10^{18} cm^{-3}, but the electron density in the conduction band reached only $n_0 \sim 10^{13}$–10^{15} cm^{-3}; the activation energy of the conduction process varied from sample to sample within the range 0.05–0.03 eV. An analysis of the electron structure of heavily doped and compensated CdTe crystals demonstrated[8] that the compensation was due to an intrinsic compensating defect in the form of the doubly charged cadmium vacancy V_{Ca}^{2+} with a level at $\varepsilon_a \sim E_c - 0.6$ eV. The application of strong electric field pulses of 10^3–10^4 V/cm intensity resulted, after a short delay t_d from the beginning of a pulse, in a considerable reduction in the voltage across a sample (which occurred in a very short time of $< 10^{-7}$ s) accompanied by a rise of the current. The delay time t_d decreased rapidly on increase in the field (values of t_d right down to 10^{-6} s were observed). A typical current-voltage characteristic is shown in Fig. 1.

The authors of Ref. 6 studied the dependence of the switching effect on the thickness d of a sample, on its temperature T, and on the rate of heat removal. It was found that the dependences of the switching field on d and T began to appear only for thick samples ($d > 300$ μm) with poor removal of heat when the pulses were relatively long ($> 10^{-3}$ s). However, in strong fields when the delay time approached $t_d \to 0$, the switching phenomena were unaffected by d or T.

The absence of the thickness dependence (at least in strong fields) and the fact that the contacts did not influence the switching led the authors of Ref. 6 to conclude that the observed effects were not due to double injection from the contacts (such injection is frequently used to account for the S-shaped characteristics).[9]

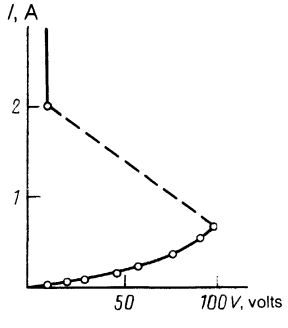

FIG. 1. Typical current-voltage characteristic of CdTe:Cl crystals, recorded at 300 K using voltage pulses of ~ 2 μs duration; the thickness of the sample was ~ 100 μm (Ref. 6).

 0038-5646/91/030515-09$03.00

As for a possible influence of the simple heating of a sample, estimates indicate that for a mobility of $\mu = 10^3$ $cm^2 \cdot V^{-1} \cdot s^{-1}$ and a density of $n = 10^{14}$ cm^{-3} (corresponding to a conductivity of $\sigma \sim 10^{-2}$ $\Omega^{-1} \cdot cm^{-1}$) in fields of $E \sim 10^4$ V/cm the heat P released per unit volume corresponded to a power density $\sim 10^6$ W/cm^3 (a direct estimate from the current-voltage characteristic in Fig. 1 for contacts of ~ 0.01 cm^2 area also gave $P \sim 10^6$ W/cm^3). Clearly, this power density was insufficient for significant heating of a sample in ~ 1 μs. On the other hand, an S-shaped current-voltage characteristic was obtained earlier[5] at a comparable rate of heat evolution when the pulses were of ~ 10 ns duration.

This circumstance, together with the absence of a strong temperature dependence of the resistance, indicated that the observed effect could not be due to a purely thermal mechanism.

A considerable rise of the conductivity (Fig. 1) could only be due to, as pointed out in Ref. 6, the ionization of some deep level (because the shallow-donor levels would have been practically completely ionized at $T \sim 300$ K). The highest density of states was associated with the presence of a compensating defect level at ε_a [in the case of the interband breakdown considered as a possibility in Refs. 4 and 5, the density of states (as deduced in a discussion below) should be less than the density of states of the compensating level ε_a]. Since by itself the process of impact (collisional) ionization in the presence of just one level could not give rise to an S-shaped current-voltage characteristic, it was suggested in Ref. 8 that the ionization occurred in a three-level system (which may be characterized by S-like behavior).[10] However, this hypothesis did not fit the real experimental situation in heavily doped compensated CdTe (Ref. 8), because even after an allowance for the large-scale potential relief (with an activation energy of the percolation level ~ 30–50 meV)[6] the shallow levels should be largely ionized at $T \sim 300$ K and there were no deep levels with a significant density of states in the interval between E_c and ε_a.

We can thus see that, in spite of systematic experimental investigations of the electronic properties of CdTe reported so far, the nature of the reversible switching effect observed at room temperature remains unclear. We shall propose a purely electronic mechanism which can account for the experimental results if we postulate a combination of the runaway effect in a strong electric field in the case of polar optical scattering, impact ionization of a deep compensating level, and electron–electron scattering. It is the electron–electron scattering that can maintain a low-resistance state in weak fields.

We shall now consider in greater detail the physical situation. Since in the polar optical scattering case the mean free path l of electrons increases on increase in their energy ε, $l \propto \varepsilon$, in a sufficiently strong electric field an electron with a sufficiently high energy acquires the energy faster from the field than it loses it to optical phonons. This results in a strong heating of the electron subsystem, giving rise to the runaway effect. The distribution function of electrons in the range of energies exceeding a certain value ε_2 (we shall show later that $\varepsilon_2 \propto E^{-2}$) is constant: $f(\varepsilon)_{\varepsilon > \varepsilon_2} = \text{const}$. Experiments indicated in the case of CdTe the polar optical scattering predominates at temperatures $T > 300$ K.

A strong influence of the runaway effect on the behavior of CdTe in strong fields was pointed out in Refs. 4 and 5. The runaway of electrons made it possible to observe the Gunn effect in pure samples using relatively weak fields ($E \sim 12$ kV/cm),[7] although the side valleys were located at $\varepsilon_L \sim 1.6$ eV above the bottom of the conduction band. On the other hand, observation of the Gunn effect indicated that the scattering to the side valleys was indeed the mechanism that limited the runaway and determined the "limiting" energy $\varepsilon = \varepsilon_L$. A detailed numerical modeling of the runaway effect in CdTe, carried out allowing for the various scattering mechanisms, was reported in Ref. 11; it was found there that a field of just $E \sim 12$ kV/cm was sufficient to transfer at least half of all the electrons to the high energy range $\varepsilon \sim \varepsilon_L$.

We could therefore expect that in a field of $E > 10^4$ V/cm the runaway effect would ensure a sufficiently high density of high-energy electrons $n_>$ with an energy $\varepsilon > \varepsilon_a$ capable of ionizing the compensation level and causing the impurity breakdown. Moreover, the relationship between $n_>$ and the value of E should be single-valued, so that the dependence of the total electron density on E should be also single-valued.

The necessary single-valued behavior could be ensured by electron–electron collisions contributing a nonlinearity to the kinetic (transport) equation of the distribution function $f(\varepsilon)$. This inhomogeneity would be easiest to understand in the limit $\tau_{ee} \to 0$ (τ_{ee} is the electron–electron relaxation time) when the electron–electron processes control the form of $f(\varepsilon)$ and determine the electron temperature T_e, which may differ from the lattice temperature. In this case if the field is sufficiently strong so that the runaway effect exists ($\varepsilon_2 < \varepsilon_L$), we can expect two stable states: 1) the bulk of electrons is concentrated in the region $\varepsilon \sim T$, where the mobility and consequently the energy acquired from the field is low: $T_{e1} \sim T$; 2) the bulk of the electrons is concentrated in the runaway region so that $T_{e2} \sim \varepsilon_L \gg T$. In the corresponding range of energies (in the main valley) the mobility and, consequently, the energy acquired from the field are high. This circumstance and the "drag" by particles from the low-mobility region ($\varepsilon \sim T$) by the high-energy electrons can ensure maintenance of the values of the temperatures $T_e \gg T$.

The possibility of a many-valued electron distribution under the conditions of the runaway effect in the presence of electron–electron collisions was pointed out by Levinson[12] in the specific case of low-temperature behavior of InSb (where the effect was somewhat different). In fact, he considered an overheating instability which was analyzed in its general form in the review of Volkov and Kogan.[13]

However, the electron–electron collisions could control completely the form of $f(\varepsilon)$ only when their density was high. This could be achieved for almost complete ionization of the deep centers ($n \sim N \sim 10^{17}$–10^{18} cm^{-3}), but it was not true of the initial situation ($n \sim n_0 \sim 10^{13}$–10^{15} cm^{-3}). Estimates of the differential conductance corresponding to the lower branch of the current-voltage characteristic in Fig. 1 indicated that n increased by no more than an order of magnitude because of the ionization corresponding to this branch. Therefore, the electron–electron processes could not be allowed for by a simple introduction of the electron temperature,[13] but would require special analysis.

In particular, we shall see later that a special role is played by the processes in which a high-energy electron from

the runaway region and characterized by $\varepsilon \sim \varepsilon_L$ is scattered by a low-energy electron with $\varepsilon' \sim T$ and transfers the latter to the runaway region where $\varepsilon > \varepsilon_2$. These processes involve a relatively small energy transfer $\varepsilon_2 \ll \varepsilon \sim \varepsilon_L$; their probability is not characterized by the transport time τ_{ee}, but by an "escape" time τ_e, which is considerably less than τ_{ee} (the reduction is in the ratio $\varepsilon_L/\varepsilon_2$, which is accurate to within a logarithmic factor).

The arrival of an electron in the runaway region (where the distribution is controlled by the electric field) due to such processes is obviously proportional to n_2/τ_e (where n_2 is the total density of electrons with energies $\varepsilon \sim \varepsilon_L$ in the runaway region). On the other hand, the escape to the low-energy region because of the electron–electron processes should be proportional to n_2/τ_{ee} (since it occurs only in the energy range $\varepsilon \sim \varepsilon_2$) and if $\varepsilon_2 \ll \varepsilon_L$ this escape effect is small compared with the arrival rate. On the other hand, it is obvious that $1/\tau_e \propto 1/\tau_{ee} \propto n_1$, where n_1 is the density of electrons with the energies $\varepsilon \sim T$, whereas the escape from the runaway region because of the electron–phonon processes is independent of n_1. Therefore, there is some critical value $n_1 = n_c$ such that if $n_1 > n_c$ the arrival of electrons in the runaway region due to the electron–electron processes cannot be compensated by the escape due to the emission of optical phonons; the density of high-energy electrons then rises to values $n_2 = n - n_c$ (we are assuming that $n_1 + n_2 = n$). The critical value of n_c depends on the electric field: if $\varepsilon_2 \ll \varepsilon_L$, we have $n_c \propto \tau_e \propto \varepsilon_L/\varepsilon_2 \propto E^{-2}$. The rise of n_c with decrease of E occurs right up to a certain critical value E_{c1} such that $1/\tau_e(E) = 1/\tau_{ee}$ (i.e., when the total flux over the spectrum from the low-energy region to the runaway region due to the electron–electron processes vanishes); we then have

$$n_c|_{E \to E_{c1}} \to \infty.$$

We shall show that in the lower branch of the current-voltage characteristic the value of n rises with the field because of the ionization processes, beginning from $n = n_0$, and that in some field E_{c2} it reaches $n(E_{c2}) = n_c(E_{c2})$, and this is followed by switching to the falling part of the current-voltage characteristic, where $n(E) = n_c(E) \propto E^{-2}$. The switching to the upper branch occurs in the field $E = E_{c1}$, which acts as the lower critical field. The behavior of the system corresponding to the upper branch is complicated by the presence of the low-mobility side valleys to which electrons scattered on increase in the field. Therefore, in strong fields $E \gg E_{c1}$ the upper branch is N-shaped and the complete current-voltage characteristic should thus demonstrate the S–N-like behavior.

1. IONIZATION PROCESSES

In our system the change in the electron density n because of the impact ionization of the deep level, considered using the simplest model situation, can be described by the equation

$$\frac{dn}{dt} = A_T(N+n_0-n) - B_T n(n-n_0) + A_I n_>(N+n_0-n), \quad (1)$$

where

$$n_> = \int_{\varepsilon_a}^{\varepsilon_L} \nu(\varepsilon) f(\varepsilon)\, d\varepsilon$$

[A_T and B_T are the thermal ionization and recombination coefficients, A_I is the impact ionization coefficient, and $\nu(\varepsilon)$ is the density of states; estimates indicate that if $n \sim 10^{18}$ cm^{-3} the Auger recombination process can be ignored]. For simplicity, we shall confine ourselves to the hydrodynamic description without analyzing the energy dependences of the various coefficients in the case of a nonequilibrium electron distribution (on the assumption that the values of these coefficients are governed by certain characteristic energies; in particular, the coefficient A_1 is determined by the contribution of electrons with energies $\varepsilon > \varepsilon_a$).

We shall also assume that the ionization processes are much slower than the establishment of an equilibrium in the electron system, so that the kinetic equation for $f(\varepsilon)$ can be solved by assuming that n is given. (It follows from the analysis given below that this is justified, at least when the ionization cross section is less than the transport cross section of the electron–electron scattering when the energy is $\varepsilon \sim \varepsilon_L$.) This in turn means that the relationship between $n_>$ and n can be regarded as given by Eq. (1).

In the steady-state case ($dn/dt = 0$) if $A_I n_> < A_T$, we have

$$n = \max(n_0,\ (A_T N/B_T)^{1/2}), \quad (2)$$

whereas in the case of predominance of the impact ionization process ($A_I n_> > A_T$) and if $n \gg n_0$, Eq. (1) reduces to

$$B_T n^2 - A_I n_> (N-n) = 0. \quad (3)$$

2. RUNAWAY EFFECT

We shall estimate the density of high-energy electrons $n_>$ by turning to the kinetic equation for the part of the distribution function $f(\varepsilon)$ symmetric in respect of the momentum. We shall assume that the most important scattering mechanism is that involving polar optical phonons (which represents the case of CdTe and $T \approx 300$ K),[11] when the relaxation time is proportional to $\varepsilon^{1/2}$ (see, for example, Ref. 14). This scattering can obviously be regarded as quasi-elastic if $\varepsilon \gg \hbar\omega_0$ ($\hbar\omega_0$ is the energy of an optical phonon, which in the case of CdTe is $\hbar\omega_0 \approx 240$ K). The process of energy relaxation is then slower than the relaxation of the momentum (see, for example, Ref. 14) and the distribution function is almost isotropic. In moderately strong fields we can assume that the energy acquired in the distance equal to the mean free path is small: $eEl(\varepsilon) \ll \varepsilon$. These circumstances allow us to regard the evolution of the electron distribution as energy diffusion (by analogy with Ref. 15) with the diffusion coefficient

$$(\hbar\omega_0)^2 \frac{1}{\tau_p}\frac{T}{\hbar\omega} + [eEl(\varepsilon)]^2 \frac{1}{\tau_p}\frac{T}{\hbar\omega}\, k_e^{-1}. \quad (4)$$

The first term allows here for the energy diffusion process associated with the stimulated electron–phonon processes and τ_p is the "escape" relaxation time in the case of the spontaneous processes; the quantity $(T/\tau_p\hbar\omega_0)k_e^{-1}$ is the reciprocal of the momentum relaxation time and the factor $k_e \approx \ln(\varepsilon/\hbar\omega_0)$ allows for the difference between the transport and "escape" relaxation times.[14] We shall adopt the approximation $T \gg \hbar\omega_0$; strictly speaking, if $T \sim 300$ K, this strong inequality is not obeyed, but because we shall be always interested in the range of high energies $\varepsilon \gg T$, such an approximation is quite reasonable.

If we allow for the spontaneous electron–phonon processes and also for the electron–electron collisions, we find that the kinetic equation satisfying the requirement that the total flux of electrons of the spectrum j_ε should vanish, becomes

$$j_\varepsilon=-\nu(\varepsilon)\left[D_\varepsilon\frac{\partial f}{\partial\varepsilon}+\frac{\hbar\omega_0}{\tau_p}f\right]+j_{ee}=0, \tag{5}$$

where j_{ee} is a flux along the spectrum due to the electron–electron collisions; the specific form of this flux will not be considered. We can see from Eq. (4) that right up to an energy

$$\varepsilon_1\sim\frac{T\hbar\omega_0 k_e^{1/2}}{eEl_0}$$

[$l_0\equiv l(\varepsilon=T)$] the expression for D_ε is dominated by the first term and if $j_{ee}=0$, we have $f(\varepsilon)\propto e^{-\varepsilon/T}$, whereas for $\varepsilon>\varepsilon_1$ we obtain $D_\varepsilon\propto\varepsilon^2/\tau_p$. If $\varepsilon>\varepsilon_1$, the formal solution of Eq. (5) can be written in the form

$$f(\varepsilon)=C\exp\left[-\frac{\varepsilon_1}{T}+\left(\frac{\varepsilon_2}{\varepsilon}-\frac{\varepsilon_2}{\varepsilon_1}\right)\right]+\int_0^\varepsilon\frac{j_{ee}(\varepsilon')}{D_\varepsilon\nu(\varepsilon')}\exp\left(\frac{\varepsilon_2}{\varepsilon'}-\frac{\varepsilon_2}{\varepsilon_1}\right)d\varepsilon', \tag{6}$$

$$\varepsilon_2=\frac{T(\hbar\omega_0)^2k_e}{(eEl_0)^2}. \tag{6a}$$

Naturally, Eq. (6) represents in practice an integral equation because the flux j_{ee} obviously itself depends on f. The simplest situation corresponds to $j_{ee}\to 0$. We then have $C=f(\varepsilon=0)$; if $\varepsilon_1<\varepsilon<\varepsilon_2$, a nonexponential fall of $f(\varepsilon)$ occurs, whereas for $\varepsilon>\varepsilon_2$ we obtain $f(\varepsilon)\approx$const (runaway effect); the "tail" of the distribution is limited by the intervalley processes at $\varepsilon\sim\varepsilon_L$ and we have

$$f(\varepsilon>\varepsilon_2)\sim f(\varepsilon_L)\sim f(0)\exp\left[-\left(\frac{\varepsilon_1}{T}+\frac{\varepsilon_2}{\varepsilon_1}\right)\right]\sim f(0)\exp\left(-2\frac{\varepsilon_1}{T}\right)$$
$$=f(0)\exp[-2\hbar\omega_0k_e^{1/2}/eEl_0].$$

Estimates for CdTe, which agree with the results of machine calculation,[11] indicate that if $E\sim 1.2\times10^4$ V/cm, then half the electrons are in the runaway tail. However, as pointed out in the Introduction, the relationship between $f(\varepsilon\sim\varepsilon_L)$ and, therefore between the ratio $n_>/n$ and the field E, is single-valued and it cannot account for the S-like behavior.

3. ELECTRON TEMPERATURE APPROXIMATION

We shall need later a more detailed knowledge of the distribution of electrons between the main and the side valleys. Obviously, the distribution function in the main valley does not terminate abruptly at $\varepsilon=\varepsilon_L$, but is governed by the occupancy of the side valleys $\Delta\varepsilon_L$ (measured from the side-valley energy ε_L). The value of $\Delta\varepsilon_L$ depends on the rate of the electron–phonon intervalley transitions and on the losses to the side valleys; from now on we shall assume that the influence of the electric field on electrons in the side valleys can be ignored. Obviously, $\Delta\varepsilon_L$ cannot exceed the depth of the side valleys ε_{Lv}. Then in the case of CdTe we have $\varepsilon_{Lv}\ll\varepsilon_L$ (we shall use this inequality in subsequent calculations).

The electron–electron processes are easiest to allow for in the situation when $\tau_{ee}\to 0$, because then these processes determine the electron distribution (the upper limit to τ_{ee} will be considered in detail later). In this case we have $f(\varepsilon)=f(0)\exp(-\varepsilon/T_e)$, where T_e is some electron temperature. In the case of the side valleys the distribution function is

$$f_v\propto\exp\left(-\frac{\varepsilon_L+\varepsilon'}{T}\right),$$

where $\varepsilon'=\varepsilon-\varepsilon_L$.

It should be noted that the condition of smallness of τ_{ee} is clearly the most stringent in the range $\varepsilon>\varepsilon_L$, where the efficiency of the electron–electron processes rises because of an increase in the phase volume ("inclusion" of the side valleys); the electron–electron exchange of energies between the main and side valleys is hindered because of the considerable difference between the effective masses. However, since the variation of $f(\varepsilon)$ in the range $\varepsilon>\varepsilon_L$ occurs on a scale $<\varepsilon_{Lv}\ll\varepsilon_L$, it follows from the results of the subsequent analysis that the contribution of the electron–electron processes to the electron flux over the spectrum rises in the range $\varepsilon>\varepsilon_L$ (for a given value of τ_{ee}) and the rise is by a factor $\sim\varepsilon_L/\Delta\varepsilon_L$, which relaxes the requirements that τ_{ee} has to satisfy. On the other hand, even if in the range $\varepsilon<\varepsilon_L$ the electron–phonon processes predominate over the electron–electron processes (see Appendix 1) in the situation when the electron–electron processes predominate if $\varepsilon>\varepsilon_L$, the electron temperature model makes it possible to describe correctly such details as a redistribution of electrons between the side valleys.

The electron temperature T_e can be found from the energy balance equation:[1)]

$$\int_0^{\varepsilon_L}\nu(\varepsilon)\left[D_\varepsilon\frac{\partial f}{\partial\varepsilon}-\frac{\hbar\omega_0}{\tau_p}f\right]d\varepsilon'-n_v\frac{\hbar\omega_0}{\tau_{pv}}=0. \tag{7}$$

Here the first term describes the energy balance in the main valley and the second term describes the losses to the side valleys; $n_v\approx f_v\nu_v$, ν_v, and τ_{pv} are, respectively, the density of electrons, the density of states, and the electron–phonon scattering time for the side valleys. We shall ignore the difference between the upper limit of integration with respect to ε' and the value of ε_L. We shall also bear in mind that the presence of the side valleys is important only if $T_e\sim\varepsilon_L$, so that we can assume that $T_e\gg\varepsilon_{Lv}$, i.e., that $f_v(\varepsilon')\approx$const. For this reason we have $f_v\sim f(\varepsilon_L)$ and $n_v\sim n\nu_v/\nu(\varepsilon_L)$ (n is the density in the main valley).

Here and later we shall describe the main valley by the model of an isotropic parabolic spectrum, because it seems to us that an allowance for the nonparabolicity or anisotropy cannot alter qualitatively the investigated phenomena (as is confirmed, in particular, by an analysis of the runaway effect for the real spectrum of CdTe reported in Ref. 11). Substituting $f\propto\exp(-\varepsilon/T_e)$ in Eq. (7), we obtain

$$\exp(-\varepsilon_L/T_e)=\frac{1-\varepsilon_2/2T_e+\varepsilon_2/2T_e}{1+\varepsilon_L/T_e-\varepsilon_2/2T_e+\varepsilon_L^2/2T_e^2+a(\varepsilon_L\varepsilon_2/2T_e^2)}, \tag{8}$$

where $a=\varepsilon_{Lv}\nu_v\tau_p(\varepsilon_L)/\varepsilon_L\nu(\varepsilon_L)\tau_{pv}(\varepsilon_{Lv})$. We can see that if $T\ll\varepsilon_2\ll\varepsilon_L$, then Eq. (8) has three roots: $T_{e1}\sim T$, $T_{e2}\sim\varepsilon_2/2$ (these roots correspond to vanishing of the numerator), and $T_{es}\sim\varepsilon_L$.

If $\varepsilon_2\sim T$, then

$$T_{e1,2}=\frac{\varepsilon_2}{4}\pm\left(\frac{\varepsilon_2^2}{16}-T\frac{\varepsilon_2}{2}\right). \quad (9)$$

The temperature T_{e3} in the $a<\varepsilon_L/\varepsilon_2$ case is described by $T_{es}\sim\varepsilon_L^2/\varepsilon_2 a$ where $n_v/n\sim a$, whereas for $a>\varepsilon_L/\varepsilon_2$, the energy T_{e3} is given by the equation

$$T_{e3}\sim\varepsilon_L/\ln\left[a\frac{\varepsilon_L\varepsilon_2}{2T_{e3}^2}\right],\quad \frac{n_v}{n}\sim\frac{2T_e^2}{\varepsilon_{Lv}\varepsilon_2}\frac{\tau_{pv}}{\tau_p}. \quad (10)$$

We can easily see that the root T_{e2} represents a solution which is unstable for a given value of ε_2, i.e., in a given field E. In fact, if we write down schematically the energy balance equation (7) in the form

$$\mathcal{F}(E,\ T_e)=Q(E,\ T_e)-P(E,\ T_e)=0, \quad (11)$$

where Q is the power acquired from the electric field [represented by the term proportional to D_ε in Eq. (7)] and P is the loss power, we can see that for $T_e=T_{e2}$ we have

$$\left.\frac{\partial\mathcal{F}}{\partial T_e}\right|_{E=\text{const}}>0,$$

which corresponds to the overheating instability of Ref. 13.

It is clear from Eq. (9) that the roots T_{e1} and T_{e2} exist only if $\varepsilon_2>8T$, which gives us an estimate of the upper critical field E_{c2}. In turn, in the limit $\varepsilon_2\to\infty$ (i.e., when $E\to 0$), we are left only with the root $T_{c1}\sim T$. The value of the lower critical field E_c is then found from the condition $T_{e2}=T_{e3}$ or, which is the same, from the condition

$$\frac{\partial\mathcal{F}(E,T_e)}{\partial T_e}=0. \quad (12)$$

It therefore follows that as in the case considered in Ref. 12, a combination of the runaway effects and of the electron–electron interaction can ensure that the current-voltage characteristic is S-shaped even in the absence of any ionization effects. Obviously, the transition or switching from the first branch ($T_e\sim T$) to the third ($T_e\sim\varepsilon_L$) results in a steep rise of the effective ionization rate and, consequently, increases the conductivity; following Eq. (3), we can assume that $n\sim N$.

A special feature of the situation considered here (as compared with the general overheating instability of Ref. 13) is the role of the intervalley processes. It follows from Eq. (10) that in the case of the upper branch of the current-voltage characteristic if $\varepsilon_2\ll\varepsilon_L$, we have $T_e\sim\varepsilon_L$ and also $n_v/n\propto 1/\varepsilon_2$, so that for a fixed value of n the number of electrons in the main valley is $(n-n_v)\propto\varepsilon_2\propto 1/E^2$; since we are ignoring the mobility in the side valleys, the conductivity decreases on increase in E.

In other words, although for $E=E_{c1}+0$ the differential conductivity is positive [$\sigma_d=dj/dE=\sigma(T_e)+E(\partial\sigma/\partial T_e)(\partial T_e/dE)$, where σ is governed by the condition $Q=\sigma E^2$; if $E=E_{c1}$, it then follows from Eqs. (11) and (12) that $dT_e/dE=\infty$ and $d\sigma/\partial T_e>0$], whereas for $E\gg E_{c1}$, we have $\sigma_d<0$. It means that the upper branch also has a falling region corresponding to the Gunn effect and the whole current-voltage characteristic is S–N shaped.

4. ANALYSIS OF THE ROLE OF THE ELECTRON–ELECTRON CONDITIONS. DISCUSSION OF THE REAL SITUATION CORRESPONDING TO THE LOWER BRANCHES OF THE CURRENT-VOLTAGE CHARACTERISTIC

As pointed out in the Introduction, although in the situation of effective ionization of the deep centers (which is reached when $T_e\sim\varepsilon_L$), we have $n\sim N\sim 10^{18}$ cm^{-3} and the electron–electron collisions may control the energy relaxation process, this is obviously not true initially ($n\sim n_0\sim 10^{13}$–10^{15} cm^{-3}). We shall therefore consider in greater detail the limitations imposed on τ_{ee} and the role of the electron–electron processes.

It is known that the influence of these processes on the momentum (or energy) relaxation can be described by the Landau collision integral (see, for example, Ref. 16). The corresponding electron flux over the spectrum considered ignoring the anisotropy can be written in the form

$$j_{ee}=\nu(\varepsilon)\varepsilon\int d\varepsilon'\,\nu(\varepsilon')\left[f(\varepsilon)\frac{\partial f}{\partial\varepsilon'}-f(\varepsilon')\frac{\partial f}{\partial\varepsilon}\right]2\varepsilon'\sigma_t|v-v'|. \quad (13)$$

Here, v is the electron velocity, σ_t is the transport cross section for the Coulomb scattering {$\sigma_t\sim\Lambda\pi e^4/[\max(\varepsilon,\varepsilon')]^2$, where $\Lambda=\ln(|p-p'|/\hbar\varkappa)$ is the Coulomb logarithm, and $\varkappa$ is the reciprocal of the screening radius of the system}; σ_t is related to the electron energy relaxation time τ_{ee} (equal to the transport momentum relaxation time) by the expression $\tau_{ee}^{-1}\propto\sigma_t nv$.

If we substitute j_{ee} from Eq. (13) into Eq. (5) and find the solution of Eq. (5) in the form $f=f(0)\exp(-\varepsilon/T_e)+f_1$, we can show that in the case of relaxation of the solution with $T_e\sim\varepsilon_L$ and $f_1\ll f$, we must satisfy the condition

$$\frac{\varepsilon_L}{\tau_{ee}(\varepsilon_L)}>\frac{\hbar\omega_0}{\tau_p(\varepsilon)}. \quad (14)$$

However, we shall draw attention to the fact that in the derivation of the expression for the Landau collision integral[14] it was essentially assumed that the problem is described by a single energy scale and that the energy transferred in a collision event is small compared with this scale. This makes it possible to carry out an expansion in terms of that small energy transfer which is responsible for the differential form of Eq. (13). However, in the situation discussed by us the runaway effect has in the absence of the electron–electron collisions three characteristic scales: 1) $\varepsilon\sim\varepsilon_L$, which represents the runaway effect where $f\sim$const; 2) $\varepsilon\sim\varepsilon_2$, which represents demarcation of the region of fall of $f(\varepsilon)$ from the runaway region; 3) $\varepsilon\sim T$, which is the region where the bulk of electrons is concentrated when $\varepsilon_2\gg T$. In this connection one should mention the processes in which an electron from the runaway region with a characteristic energy $\varepsilon'\sim\varepsilon_L$ is scattered by a low-energy electron with an energy $\varepsilon''\sim T$, and transfers the latter to the runaway region $\varepsilon>\varepsilon_2$, where the distribution is controlled by the electric field. In describing these processes the transfer energy $\Delta\varepsilon$ cannot be in general regarded as small, because $\Delta\varepsilon>\varepsilon_2>\varepsilon'$.

We can show (see the Appendix 2) that the contribution of such processes to j_{ee} when $\varepsilon\sim\varepsilon_2$ is of the order of

$$\tilde{j}_{ee}\propto\int_0 d\varepsilon''\,\nu(\varepsilon'')f(\varepsilon'')\int_{\varepsilon'>\varepsilon}d\varepsilon'\,\nu(\varepsilon')\tilde{\sigma}(\varepsilon')v'\frac{\varepsilon'}{\varepsilon}f(\varepsilon'), \quad (15)$$

where $\tilde{\sigma}$ differs from σ_t by the absence of the logarithmic factor Λ. We shall allow for the fact that since for $\varepsilon>\varepsilon_2$, we have $f(\varepsilon')\sim\text{const}\sim f(\varepsilon\sim\varepsilon_L)\equiv f_L$ and for $\varepsilon<\varepsilon_2$ the function $f(\varepsilon)$ falls rapidly on increase in ε, we obtain, denoting by n_1 the density of electrons in the range $\varepsilon<\varepsilon_2$:

$$\tilde{j}_{ee} \sim \frac{\varepsilon_L}{\varepsilon} \tilde{\sigma}(\varepsilon_L) v(\varepsilon_L) \varepsilon_L f_L n_1 = \frac{\varepsilon_L}{\varepsilon \tilde{\tau}_{ee}(\varepsilon_L)} f_L \nu(\varepsilon_L) \varepsilon_L, \tag{16}$$

where $\tilde{\tau}_{ee}^{-1} \equiv \tilde{\sigma} v n_1$. It should be noted that if $\varepsilon_2 \ll \varepsilon_L$, this contribution exceeds by a factor $\sim(\varepsilon_L/\varepsilon_2\Lambda)$ an estimate obtained from Eq. (13). This is due to the fact that the contribution of "low-angle" processes ($\Delta\varepsilon \ll \varepsilon''$, ε') j_{ee}^{L} is governed by the transport relaxation time τ_{ee}, whereas $\tilde{j}_{ee}$ is governed by the "escape" time, i.e., by the probability of a single scattering event with a finite energy transfer. In the case of the Coulomb scattering the latter diverges for small transferred momenta q since $q^{-2} \propto (\Delta\varepsilon)^{-1} \propto \varepsilon_2^{-1}$.

We shall now write down Eq. (5) allowing both for j_{ee}^{L} and for $\tilde{j}_{ee}$ in the case when $n_1 \sim n$ (i.e., in the situation corresponding to the lower branch of the current-voltage characteristic when the bulk of electrons is concentrated at low energies) and $\varepsilon \sim \varepsilon_2$:

$$-\nu(\varepsilon) D_\varepsilon \frac{\partial f}{\partial \varepsilon} + \frac{\varepsilon_L}{\varepsilon} \frac{\varepsilon_L \nu(\varepsilon_L)}{\tilde{\tau}_{ee}(\varepsilon_L)} f_L = \frac{\varepsilon}{\tau_{ee}(\varepsilon)} \nu(\varepsilon) f(\varepsilon) + \frac{\hbar\omega_0}{\tau_p(\varepsilon)} \nu(\varepsilon) f, \tag{17}$$

where $\tau_{ee}^{-1}(\varepsilon) = \sigma_t(\varepsilon) v(\varepsilon) n_1$. We are allowing here for the fact that if $n_1 \sim n$ and $\varepsilon_2 \gg T$, the second term in Eq. (13) can be ignored relative to the ratio T/ε_2. A comparison of Eqs. (17) and (5) shows that, since low-angle electron–electron scattering, together with the electron–phonon processes, results in the escape of electrons from the runaway region, the processes characterized by a finite energy transfer "assist" the formation of the runaway region by the electric field.

We shall use, as before, a parabolic model of the electron spectrum (which may be justified at least when $\varepsilon_2 \ll \varepsilon_L$) and we note that then $\varepsilon/\tau_{ee} \propto \varepsilon^{-1/2} \propto \tau_p^{-1}$. Bearing this in mind, we find [by analogy with Eq. (6)] that

$$f \approx \int_{\varepsilon}^{} d\varepsilon' \frac{\nu(\varepsilon_L) \varepsilon_L f_L}{\tilde{\tau}_{ee}(\varepsilon_L)} \frac{\varepsilon_L}{\varepsilon'} \frac{1}{D_\varepsilon(\varepsilon') \nu(\varepsilon')} \exp\left(\frac{\tilde{\varepsilon}_2}{\varepsilon} - \frac{\tilde{\varepsilon}_2}{\varepsilon'}\right) + C \exp\left[-\frac{\varepsilon_1}{T} - \left(\frac{\tilde{\varepsilon}_2}{\varepsilon} - \frac{\tilde{\varepsilon}_2}{\varepsilon_1}\right)\right], \tag{18}$$

where $\tilde{\varepsilon}_2$ is determined allowing for the small-angle electron–electron processes:

$$\tilde{\varepsilon}_2 \sim \frac{T^2 \tau_p}{(eEl_0)^2} \left(\frac{\hbar\omega_0}{\tau_p} + \frac{\varepsilon}{\tau_{ee}(\varepsilon)}\right) \sim \varepsilon_2 \left(1 + \frac{\varepsilon \tau_p}{\tau_{ee}(\varepsilon) \hbar\omega_0}\right). \tag{19}$$

For low values of n the difference between $\tilde{\varepsilon}_2$ and ε_2 can obviously be ignored. If $\varepsilon < \tilde{\varepsilon}_2$, we can easily see that the first term in Eq. (18) can be ignored compared with the second. If $\varepsilon > \tilde{\varepsilon}_2$, then bearing in mind that $f(\varepsilon) \sim f_L$ in Eq. (18) and using the normalization condition

$$C \sim n/\nu(T) T, \tag{20}$$

we obtain

$$f_L \sim n_1/\nu(T) T \left[1 - \frac{\varepsilon_L}{\tilde{\tau}_{ee}(\varepsilon_L)} \frac{\tau_p(\varepsilon_L)}{\hbar\omega_0 k_e} \frac{\varepsilon_L}{\tilde{\varepsilon}_2}\right]^{-1} \exp\left(-\frac{\varepsilon_1}{T} - \frac{\tilde{\varepsilon}_2}{\varepsilon_1}\right). \tag{21}$$

We shall first consider the situation when

$$\frac{\varepsilon}{\tau_{ee}(\varepsilon)} \ll \frac{\hbar\omega_0}{\tau_p}, \tag{22}$$

which corresponds to the lower branch of the current-voltage characteristic. We then have $\tilde{\varepsilon}_2 = \varepsilon_2$ in Eq. (21) and the argument of the exponential function is $-2\hbar k_e^{1/2}/(eEl_0)$. The total number of electrons in the runaway region $n_2 \equiv f_L \nu(\varepsilon_L) \varepsilon_L$ is then given by

$$n_2 \sim n_1 \beta \frac{n_c}{n_c - n_1}, \quad \beta = \frac{\nu(\varepsilon_L) \varepsilon_L}{\nu(T) T} \exp\left(-\frac{2\hbar\omega_0 k_e^{1/2}}{eEl_0}\right), \tag{23}$$

where

$$n_c(E) \equiv \frac{\varepsilon_2}{\varepsilon_L} N \frac{\hbar\omega_0 k_{ee}}{\tau_p(\varepsilon_L)} \frac{\tilde{\tau}_{ee}(\varepsilon_L, n=N)}{\varepsilon_L}. \tag{24}$$

It is clear from Eq. (24) that n_1 should not exceed the value n_c and that in the limit $n \to n_c$ we have $n_2 \to \infty$. This condition sets the limits of the validity of our assumption to $n_1 \sim n$ and $n_2 \ll n_1$, so that in this section we shall ignore the scattering of high-energy electrons by one another.

We can determine n_1 using the normalization condition

$$n_1 + n_2(1+\gamma) = n. \tag{25}$$

Here, the coefficient γ (generally dependent on E) allows for a redistribution of the high-energy electrons to the side valleys (see the Appendix 1); we recall that, as assumed above, the electron distribution is established faster than the ionization processes. We can easily show that Eqs. (23) and (25) lead to

$$n_2 = \begin{cases} \dfrac{n_c n}{n_c - n} \beta(1+\gamma), & \dfrac{n_c - n}{1+\gamma} < n_c \beta^{1/2}, \quad (26a) \\ n_c \beta^{1/2}, & \dfrac{|n_c - n|}{1+\gamma} < n_c \beta^{1/2}, \quad (26b) \\ (n - n_c)/(1+\gamma), & \dfrac{n - n_c}{1+\gamma} > n_c \beta^{1/2}. \quad (26c) \end{cases}$$

Obviously, this applies also in the case when the inequality (22) is not obeyed, i.e., when $\tilde{\varepsilon}_2$ differs from ε_2. Then $\tilde{\varepsilon}$ itself depends on τ_{ee} and, therefore, on n_1; this circumstance results in the replacement of Eq. (24) with

$$\varepsilon_2/\varepsilon_L \to (\varepsilon_L/\varepsilon_2 - \Lambda)^{-1}. \tag{27}$$

It therefore follows that the electron–electron processes can increase the number of electrons in the high-energy region also when the condition (14) is disobeyed. We shall now analyze the current-voltage characteristic with allowance for this situation.

It is obvious that the rise of E in the lower branch of the current-voltage characteristic is accompanied by an increase in n because of the ionization processes and, consequently, because of the conductivity $\sigma \propto n$. Bearing in mind that, on the one hand, in this situation an ionizing defect is a negatively charged center[8] and, on the other, allowing for the fact that the velocity of electrons in the side valleys is small, we shall ignore the contribution of the latter to the ionization process. Substituting $n_> = n_2$ into Eq. (3) and describing n_2 by Eq. (26a) [on the assumption that $n < n_c(E)$], we obtain

$$n \sim \frac{n_c(E)}{2} \pm \left\{\left[\frac{n_c(E)}{2}\right]^2 - \frac{A_I}{B_r} \beta N n_c(E) (1+\gamma)\right\}^{1/2}. \tag{28}$$

Next, we obtain the condition for an instability point E_{c2} where $\partial\sigma/\partial E \propto \partial n/\partial E \to \infty$:

$$n_e(E_{c2}) = 4\frac{A_I}{B_T}\beta(1+\gamma)$$

or

$$\frac{\varepsilon_2(E_{c2})}{\varepsilon_L}\frac{\tau_{ee}(\varepsilon_L, n=N)}{\tau_p(\varepsilon_L)}\frac{\hbar\omega_0 k_e}{\varepsilon_L} = 4\frac{A_I}{B_T}\beta(1+\gamma). \qquad (29)$$

The point E_{c2} corresponds to the transition or switching to the second (falling) branch of the current-voltage characteristic, which is due to reversal of the sign in Eq. (28); for this branch we have $n \sim n_c$, so that a reduction in E [leading to an increase in $n_c(E)$] is accompanied by an increase in n (and, therefore, in σ).

It follows from Eq. (26a) that Eq. (28) is valid right up to the values

$$n_e(E) \sim \frac{A_I}{B_T}(1+\gamma)N\beta^{1/2};$$

in the case of higher values of n_c, we have to describe n_2 using Eq. (26b) and then the equation for n reduces to

$$n = \alpha\left\{\frac{N+n_c}{2} \pm \left[\frac{(N+n_c)^2}{4} - \frac{n_c N}{\alpha}\right]^{1/2}\right\}, \qquad (30)$$

$$\alpha = \tilde{A}_I/(B_T+\tilde{A}_I),\ \tilde{A}_I = A_I/(1+\gamma).$$

We can easily see that in the case of this branch the derivative on the right-hand side of Eq. (1) with respect to n is positive if $n < A_1/N/B_T$ both in the case described by Eq. (26a) and for Eq. (26b). Consequently, this branch, controlled by the ionization processes, is unstable against fluctuations of n when the distribution of E is homogeneous.

In the range of weak fields this branch is limited to a certain critical field E_{c1} found from the condition

$$N+n_c(E_{c1}) = 2[n_c(E_{c1})N/\alpha]^{1/2}, \qquad (31)$$

or—when an allowance is made for Eqs. (24) and (27)—from the condition

$$\frac{\varepsilon_L}{\varepsilon_2(E_{c1})} - \Lambda \sim \frac{1}{\alpha}\frac{\hbar\omega_0}{\tau_p(\varepsilon_L)}\frac{\tilde{\tau}_{ee}(\varepsilon_L, n=N)}{\varepsilon_L} \equiv \xi. \qquad (32)$$

We can see that if $\xi < 1$, then $\varepsilon_2(E_{c1}) \approx \varepsilon_L/\Lambda$. If $E \to E_{c1}$, we obtain

$$\frac{d\sigma}{dE} \propto \frac{dn}{dE} \propto \frac{dn}{dn_e} \to \infty;$$

the transition then takes place to the rising branch, where

$$n \propto \alpha N. \qquad (33)$$

It should be noted that if $E = E_{c1}$, then $n_2 \propto B_T n^2/A_I N \propto B_T \alpha n/A_I$, i.e., the number of high-energy electrons is of the same order of magnitude as the total number of electrons. If $\xi < 1$, then the electron temperature (at least that in the main valley) is established at the upper branch of the current-voltage characteristic. We can then see that in the range of fields from E_{c1} to $2^{1/2}E_{c1}$ the density n rises (and, therefore, the conductivity σ increases) by the factor ξ^{-1}.

We have ignored here the behavior of the coefficient γ describing the redistribution of electrons between the valleys. It follows from the estimates (see the Appendix 1) that it rises on increase in E. On the other hand, the number of electrons in the main valley (which for the upper branch is of the order of n_2) in the case when $E \gg E_{c1}$ is of the order of $N\alpha/(1+\gamma) \propto (1+\gamma)^{-2}$ if we allow for Eq. (33). Therefore, following the results of Sec. 3, we find that the upper branch has a falling region whose nature is governed by details of the intervalley kinetics. The final form of the current-voltage characteristic is represented schematically in Fig. 2.

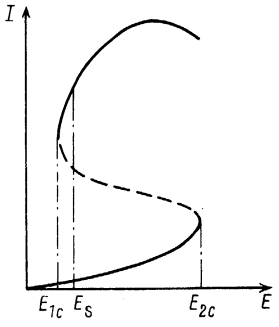

FIG. 2. Schematic representation of the current-voltage characteristic using the adopted model for the case of a homogeneous distribution of the current in a sample; $E_S \sim E_{1c}$ is the field corresponding to the formation of a filament.

5. FILAMENTATION OF THE CURRENT

If the current-voltage characteristic is S-shaped, the homogeneous distribution of the current breaks down and a filamentation-type instability is observed,[13] as a result of which a current filament appears at some value of the field $E = E_S$ in the constant-current case (which applies to the low-resistance state). The boundary of the filament can be described obviously by adding a gradient term to the kinetic equation, so that this equation becomes

$$\frac{\partial}{\partial\varepsilon} j_\varepsilon + \operatorname{div}[D_r(\varepsilon)\nabla f + \mu\nabla\varphi] = 0. \qquad (34)$$

Here, $D_r(\varepsilon)$ is the spatial diffusion coefficient, $\mu(\varepsilon)$ is the mobility, and φ is the electrical potential whose appearance is associated with local deviation from the electron density $n = \int d\varepsilon \nu(\varepsilon) f(\varepsilon)$ (in this analysis of the boundaries of the filament we shall ignore electrons in the side valleys because their relative contribution is small if the field is weak) from the concentration of ionized impurities $\tilde{n}$. Assuming that the characteristic screening length (governed primarily by the density of the low-energy electrons n_1) is less than the filament thickness [which is clearly of the order of the diffusion length of the high-energy electrons $L_D \sim D_r(\varepsilon \sim \varepsilon_L)\tau_p\varepsilon_L/\hbar\omega_0$], we find that the quasineutrality condition is $(\tilde{n} - n) \ll n$. In this case in the ionization equation we can ignore the difference between $\tilde{n}$ and n. If then $n_>$ and n in Eq. (1) are expressed in terms of the solution of Eq. (34), which corresponds to a certain potential φ, Eq. (1) can be regarded as the equation for φ which makes the problem self-consistent. The role of the potential φ then reduces to a redistribution of the low-energy electrons ensuring that the electrical neutrality condition is obeyed.

The necessary condition for the parameters in Eq. (1) is

$$\int dx \frac{\partial}{\partial\varepsilon} j_\varepsilon = 0, \qquad (35)$$

where x is the coordinate at right-angles to the boundary of the filament (the thickness of the boundary itself is assumed to be small compared with the filament diameter). The expression (35) represents the condition of stability of the filament boundary, which can be used to determine the value of E_S. To the nearest order of magnitude, this condition can

be regarded as the requirement that the spatial flux of the high-energy electrons to the "cold" region, proportional to $\sim D_r[n_2(x\to-\infty)/L_D]]$, should be compensated in the "hot" region by the upward flux along the spectrum

$$D_r[n_2(x\to-\infty)/L_D]$$
$$\sim L_D\left[D_\varepsilon/\varepsilon_L+\left(\frac{\varepsilon_L}{\varepsilon_2}-\Lambda\right)\frac{\varepsilon_L}{\tilde{\tau}_{ee}}\right]\nu(\varepsilon_L)f(\varepsilon_L)|_{x\to-\infty}, \quad (36)$$

which leads to

$$\left(\frac{\varepsilon_L}{\varepsilon_2}-1\right)\frac{\hbar\omega_0}{\tau_p}\sim\left(\frac{\varepsilon_L}{\varepsilon_2}-\Lambda\right)\frac{\varepsilon_L}{\tilde{\tau}_{ee}}. \quad (37)$$

Using Eq. (32), we then obtain

$$E_S\sim E_{c1}. \quad (38)$$

6. DISCUSSION OF RESULTS

It is clear from the above analysis that the proposed model can describe qualitatively the observed current-voltage characteristic. In the specific case of CdTe [electron mobility $\sim 10^3$ cm$^2\cdot$V$^{-1}\cdot$s^{-1}, $\tau_p(\varepsilon\sim T\sim 300$ K$)\sim 10^{-13}$ s, and $l_0\sim 300$ Å], we obtain the following order-of-magnitude estimates: $E_{c2}\sim 10^4$ V/cm and $E_{c1}\sim 10^3$ V/cm, which are again in agreement with the experimental results.

We shall now consider the problems ignored in our simplified model. We shall begin with the problem of heat release. Under the conditions of the low-resistance state in a filament when $E=E_S\sim E_{c1}$ and $n\sim 10^{17}$ cm^{-3} the power transferred to the lattice is

$$P\sim\frac{\hbar\omega_0}{\tau_p(\varepsilon_L)}n_2\sim 10^8 \text{ W/cm}^3.$$

Therefore, even in a time interval $\sim 10^{-6}$ s the temperature may rise by ~ 100 K if we ignore the removal of heat from a filament. On the other hand, it follows from Eq. (6a) that, since $l_0(T)\propto T/(T/\hbar\omega_0)=$ const, we have $\varepsilon_2\propto T$. Therefore, for a fixed value of E the rise in T increases n_c and then, in accordance with Eq. (30), the value of n falls and so does the conductivity σ (in this case in the vicinity of E_{c1} we have $\partial\sigma/\partial T\to\infty$).

These circumstances (ensuring concentration of the current in a region with lower temperatures) lead to an instability of the filament geometry (which may result in a shift of the filament to a less heated region or may cause splitting of the filament) accompanied by slight changes in the total cross section area of the filament. It follows from an analysis of the energy balance in a filament, carried out allowing for heat evolution and the phonon thermal conductivity, that if $P\sim 10^8$ W/cm^3 and $D_{\text{ph}}\sim 1$ cm^2/s (where D_{ph} is the thermal diffusivity), then for $\Delta T\approx 100$ K the evolution of heat is compensated by the removal of heat if the filament radius is $\sim 10^{-3}$ cm. On the other hand, when the total current is ~ 2 A for $c\sim 10^{17}$ cm^{-3} and $\sigma\sim 10^2$ $\Omega^{-1}\cdot$cm^{-1}, the total cross-section area of the filament $s\sim 10^{-4}$ cm^2.

It therefore follows that the mechanism under discussion here can ensure a feedback preventing a strong local heating because an increase in T reduces the mobility.

On the other hand, the thermal ionization process increases the conductivity when the temperature rises. In particular, if $\varepsilon_d\sim 0.6$ eV, $n_0\sim$ cm^{-3}, and $N\sim 10^{17}$–10^{18} cm^{-3} an increase in the temperature to ~ 600 °C should raise σ by almost two orders of magnitude. This mechanism clearly stimulates switching in relatively weak fields if the pulse duration is sufficient (large values of t_d).[6] An important factor can then be an inhomogeneity due to a large-scale potential relief (ignored in the model adopted by us), which is important for samples with the conduction activation energy $\geqslant 50$ meV. However, in this case the thermal mechanism clearly combines with that discussed above and ensures the necessary critical carrier density.

Obviously, the local heating which can cause modification of the impurity system is involved in the process of "conditioning" of p-CdTe in fields $E>2\times 10^4$ V/cm (Ref. 17); the subsequently observed reversible switching effect[17] is clearly due to the proposed electron mechanism.

It should be pointed out that although, according to our model, the nonlinearity of the lower branch of the current-voltage characteristic is entirely due to the ionization of the deep centers, such a nonlinearity may be also associated with the activation of carriers to the percolation level by an electric field.[18] Another possible source of this nonlinearity is the influence of the electric field on the thermal ionization of deep centers[19] (although it is not quite clear whether this effect can appear in fields $E<10^4$ V/cm. However, all these effects can only increase n corresponding to the lower branch of the current-voltage characteristic and, therefore, help to reach the critical density at which the switching occurs.

It is clear from Fig. 1 that at the switching or transition point the value of σ_d is finite, i.e., that the switching occurs in fields somewhat less than the critical value. Obviously, this is the reason for the existence of a delay time, which is necessary for the development of an appropriate fluctuation. The nature of such a finite fluctuation is however not yet fully understood.

We shall now identify the factors which can, in principle, help maintain the low-resistance state. One of them is the screening of the polar optical scattering under the conditions of a high electron density; however, estimates indicate that even if $n\sim 10^{18}$ cm^{-3} this effect cannot alter greatly the scattering efficiency. Another factor is a reduction in the effectiveness of the ionization processes when the centers become liberated; however, the ionization processes do not represent the main mechanism of the energy losses.

We note finally that since the width of the band gap of CdTe is $E_g\sim 1.5$ eV $<\varepsilon_L$, it follows that, in addition to the ionization of the impurity centers, we can expect also the band–band ionization (interband breakdown) which clearly accounts for the observed recombination radiation.[4] However, since $\varepsilon_L-E_g\sim 0.1$ eV $\ll E_g$, the phase volume of these processes is relatively small.

We have therefore demonstrated that the electron subsystem of doped compensated semiconductors with polar optical scattering can exhibit bistability in strong electric fields. We proposed a new purely electronic mechanism of the switching effect, which combines the runaway effect, impact ionization, and electron–electron scattering, and which accounts for the experimental data on CdTe reported in Refs. 4–6. In contrast to the purely thermal mechanism,[1] this electronic mechanism does not require the existence of a delay time and the switching time is governed by the characteristic energy relaxation times and amounts to 10^{-10}–10^{-11}

s. It seems that this mechanism can appear also in other "dirty" II-VI compounds with sufficiently high side valleys.

The authors are grateful to L. E. Vorob'ev, I. B. Levinson, M. E. Levinshteĭn, K. D. Tsendin, and I. N. Yassievich for discussing this investigation and making valuable comments, and to Yu. M. Gal'perin for reading the manuscript and making helpful suggestions.

APPENDIX 1. ANALYSIS OF A REDISTRIBUTION OF ELECTRONS BETWEEN THE VALLEYS UNDER CONDITIONS OF PREDOMINANCE OF THE ELECTRON–PHONON PROCESSES

An allowance for the transitions from the main valley to the side valleys and back again, involving the participation of optical phonons, yields the following estimate of n_v:

$$n_v = f(\varepsilon_L)\nu_v(\Delta\varepsilon)\max\left[(D_\varepsilon\tau_{0v})^{1/2}, \frac{D_\varepsilon\tau_{0v}}{\varepsilon_{Lv}}\right] \equiv \gamma f(\varepsilon_L)\nu(\varepsilon_L)\varepsilon_L, \tag{A1.1}$$

where $\Delta\varepsilon = \min(D_\varepsilon\tau_{0v}, \varepsilon_{Lv})$ and τ_{0v} is the time of the phonon-assisted transition from the main valley to the side valleys; we shall assume that $D_\varepsilon\tau_{pv} < \hbar\omega_0\varepsilon_{Lv}$, where τ_{pv} is the electron–phonon relaxation time representing the transfer to the side valleys. The density of electrons of energy $\varepsilon - \varepsilon_L \gg T$, governing the losses in the side valleys, is of the order of

$$\tilde{n}_v \sim \nu(\varepsilon_L)f(\varepsilon_L)\frac{\tau_{pv}(\Delta\varepsilon)}{\tau_{0v}(\Delta\varepsilon)\hbar\omega_0 k_e}(D_\varepsilon\tau_{0v}) \tag{A1.2}$$

and therefore the power transferred to phonons is

$$Q \sim \tilde{n}_v\frac{\hbar\omega_0 k_e}{\tau_{pv}} \sim \nu(\varepsilon_L)f(\varepsilon_L)D_\varepsilon, \tag{A1.3}$$

which exceeds by the factor $\sim\varepsilon_L/\varepsilon_2$ the losses in the main valley when $f(\varepsilon)\approx$const.

APPENDIX 2. ANALYSIS OF ELECTRON–ELECTRON PROCESSES WITH A FINITE TRANSFER OF THE MOMENTUM

We shall use a reasoning similar to that employed in the derivation of the Landau collision integral,[14] but we shall not assume that the transferred momentum q is small compared with the momentum of one of the colliding electrons (p'). We then obtain the following expression for the flux of particles across a surface $\varepsilon(p) = $ const due to collisions of electrons with the momenta p' and p'':

$$\tilde{j}_{ee}(\varepsilon) = \int_{|\mathbf{p}''|<p} d^3p'' \int d^3p' \int d^3q\, W(\mathbf{p}'', \mathbf{p}', \mathbf{q}) \times\{[f(\mathbf{p}'+\mathbf{q})-f(\mathbf{p}')]f(\mathbf{p}'') + f(\mathbf{p}')[f(\mathbf{p}'')-f(\mathbf{p}''+\mathbf{q})]\}, \tag{A2.1}$$

where W is the scattering probability. If we assume that $p' \gg p''$ and $q \ll p'$, and bear in mind that integration with respect to q is governed by the lower limit $q_{\min}$, which is the minimum transferred momentum necessary to supply an energy exceeding ε to a low-energy electron, we find that (compared with Ref. 14)

$$\int_0^\varepsilon d\varepsilon''\nu(\varepsilon'')f(\varepsilon'')\int d^3p'\,\tilde{\sigma}_t(\varepsilon')\frac{\varepsilon'^2}{q_{\min}}\frac{\partial f}{\partial\varepsilon'} - \int_0^\varepsilon d\varepsilon''\nu(\varepsilon'')[f(\varepsilon''+\Delta\varepsilon_{\min}) - f(\varepsilon'')]\int_{\varepsilon'>\varepsilon} d^3p'\,\tilde{\sigma}_t(\varepsilon')\nu'\frac{\varepsilon'}{\varepsilon}f(\varepsilon'). \tag{A2.2}$$

Since in the range $\varepsilon < \varepsilon_2$ the function $f(\varepsilon'')$ falls rapidly on increase in ε'', we can assume that it differs from zero only when $\varepsilon'' \sim T$, so that $q_{\min} \sim p$ and $\Delta\varepsilon_{\min} \sim \varepsilon$. Moreover, bearing in mind that in the runaway region we have $f(\varepsilon') \approx$ const, we can show that the first term in Eq. (A2.2) can be ignored compared with the second, which gives us the estimate described by Eq. (15).

[1)] This equation is readily obtained if the kinetic equation $\dot{f} + \mathrm{div}_\varepsilon j_\varepsilon = 0$ is multiplied by ε and integrated with respect to ε' bearing in mind that in the quasiequilibrium distribution case we have $j_{ee} = 0$.

[1] V. B. Sandomirskiĭ and A. A. Sukhanov, Zarubezh. Radioelektron. No. 9, 68 (1976).
[2] S. M. Ryvkin, Pis'ma Zh. Eksp. Teor. Fiz. **15**, 632 (1972) [JETP Lett. **15**, 448 (1972)].
[3] V. F. Bannaya, E. M. Gershenzon, and L. B. Litvak-Gorskaya, Fiz. Tekh. Poluprovodn. **2**, 978 (1968) [Sov. Phys. Semicond. **2**, 812 (1969)]; A. R. Gadzhiev, S. M. Ryvkin, and I. S. Shlimak, Pis'ma Zh. Eksp. Teor. Fiz. **15**, 605 (1972) [JETP Lett. **15**, 428 (1972)].
[4] G. S. Picus, D. F. DuBois, and L. B. Van Atta, Appl. Phys Lett. **12**, 81 (1968).
[5] M. R. Oliver, A. L. McWhorter, and A. G. Foyt, Appl. Phys. Lett. **11**, 111 (1967).
[6] N. V. Agrinskaya, M. V. Alekseenko, and O. A. Matveev, Fiz. Tekh. Poluprovodn. **9**, 519 (1975) [Sov. Phys. Semicond. **9**, 341 (1975)].
[7] A. G. Foyt and A. L. McWhorter, IEEE Trans. Electron Devices **ED-13**, 79 (1966).
[8] C. E. Jones, K. James, J. Merz, R. Braunstein *et al.*, J. Vac. Sci. Technol. A **3**, 131 (1985).
[9] A. G. Milnes, *Deep Impurities in Semiconductors*, Wiley, New York (1973).
[10] A. G. Zabrodskiĭ and I. S. Shlimak, Fiz. Tverd. Tela (Leningrad) **16**, 2350 (1974) [Sov. Phys. Solid State **16**, 1528 (1975)].
[11] G. Ottaviani, Rev. Phys. Appl. **12**, 249 (1977).
[12] I. B. Levinson, Fiz. Tverd. Tela (Leningrad) **7**, 1362 (1965) [Sov. Phys. Solid State **7**, 1098 (1965)].
[13] A. F. Volkov and Sh. M. Kogan, Usp. Fiz. Nauk **96**, 633 (1968) [Sov. Phys. Usp. **11**, 881 (1969)].
[14] K. Seeger, *Semiconductor Physics*, Springer Verlag, Berlin (1974).
[15] B. I. Davydov, Zh. Eksp. Teor. Fiz. **6**, 463 (1936).
[16] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics*, Pergamon Press, Oxford (1981).
[17] N. V. Agrinskaya, M. V. Alekseenko, and O. A. Matveev, Fiz. Tekh. Poluprovodn. **9**, 1962 (1975) [Sov. Phys. Semicond. **9**, 1286 (1975)].
[18] B. I. Shklovskiĭ, Fiz. Tekh. Poluprovodn. **13**, 93 (1979) [Sov. Phys. Semicond. **13**, 53 (1979)].
[19] V. Karpus and V. I. Perel', Zh. Eksp. Teor. Fiz. **91**, 2319 (1986) [Sov. Phys. JETP **64**, 1376 (1986)].

Translated by A. Tybulewicz