

## CONCERNING THE THEORY OF IMPACT IONIZATION IN SEMICONDUCTORS

L. V. KELDYSH

P. N. Lebedev Physics Institute, Academy of Sciences, U.S.S.R.

Submitted to JETP editor January 9, 1965

J. Exptl. Theoret. Phys. (U.S.S.R.) **48**, 1692-1707 (June, 1965)

The energy distribution of the electrons in a covalent semiconductor in the presence of a strong electric field is determined. It is shown that the number of ionizing electrons increases with increasing field  $\mathbf{E}$ , first approximately like  $\exp(-\text{const} \cdot E^{-1})$ , and in extremely strong fields like  $\exp(-\text{const} \cdot E^{-2})$ .

THE average probability of impact ionization for electrons in a strong electric field is determined essentially by the probability that the electron will acquire in the field an energy equal to the ionization threshold energy  $\epsilon_i$ . The increase in the electron energy depends on the relation between two factors: acceleration in the external field and energy dissipation by collision with phonons. From this point of view, we can visualize two extreme possibilities, as a result of which the electron would acquire an energy  $\epsilon_i$ . First, it can receive this energy from the field without experiencing accidentally even a single collision; second, the same energy can be attained gradually, after many collisions, such, however, that in each collision the electron loses on the average less energy than it receives from the field during the time between two collisions.

The first of these possibilities was pointed out already in the first papers of Townsend on discharge in gases, and was investigated theoretically by Shockley for the case of ionization in semiconductors.<sup>[1]</sup> In this case the probability of impact ionization  $w_i$  is proportional to the probability that the electron will cover without collision a path  $L = \epsilon_i/eE$ , i.e.,

$$w_i \sim \exp(-\epsilon_i/eEl), \quad (1)$$

where  $l$  is the mean free path. A formula of the type (1) was first used by Chynoweth<sup>[2]</sup> to describe the experimental data on impact ionization in semiconductors.

The second possibility was investigated by Druyvesteyn,<sup>[3]</sup> Davydov,<sup>[4]</sup> and many other authors and was applied to an investigation of the breakdown in covalent semiconductors by Wolff.<sup>[5]</sup> In this case the increase of the electron energy has the character of energy diffusion. The stationary distribution of the electrons with respect

to the energy is described by the distribution function

$$f_0(\epsilon) = \text{const} \cdot \exp\left(-\frac{\epsilon^2}{e^2 E^2 l^2} \delta\right), \quad (2)$$

where  $\delta$  is a small quantity on the order of the ratio of the energy lost by the electron in one collision to the total energy. For elastic collisions in gases,  $\delta$  is of the order of the ratio of the mass of the electron to the mass of the ion,<sup>[3]</sup> while for collisions with acoustic phonons in semiconductors<sup>[4]</sup>  $\delta \sim mc^2/kT$ , where  $c$  is the speed of sound and  $T$  the temperature. Finally, for the case of interaction between electrons and optical phonons in semiconductors<sup>[5]</sup> or with molecular vibrations in gases, we have  $\delta \sim \hbar\omega/\epsilon$ , where  $\hbar\omega$  is the energy of the emitted vibrational quantum. In the latter case, the velocity distribution of the electrons (2) becomes Maxwellian with effective temperature  $kT_e \sim (eEl)^2/\hbar\omega$ .<sup>[5]</sup>

The essential result contained in (2) is the fact that the number of the ionizing electrons, which is proportional to  $f_0(\epsilon_i)$ , increases with increasing field like

$$w_i \sim \exp(-\text{const}/E^2). \quad (3)$$

Thus, the two possibilities considered lead to essentially different dependences of the probabilities of the impact ionization on the field, (1) and (3). At first glance the result (3), derived from the kinetic equation, appears to be better founded. However, the experimental data<sup>[2,6,7]</sup> confirm the existence of a relation of the type (1) over a wide range of field intensities  $\mathbf{E}$ . This contradiction was clarified by Baraff,<sup>[8]</sup> who showed that in the region of relatively weak fields, when  $eEl \lesssim \hbar\omega$  ( $\omega$  is the average frequency of the emitted phonons), the kinetic equation also leads to a dependence of the type (1), while the diffusion approx-

imation cannot be employed for its solution. If we assume for the mean free path in semiconductors a value  $l \sim 10^{-6}$  cm, and for the phonon energies  $\hbar\omega \sim 5 \times 10^{-2}$  eV, then the field in which the diffusion approximation and its corollary (3) are valid should be much stronger than  $5 \times 10^4$  V/cm. However, the breakdown fields in germanium and silicon are of the order of  $10^5$  V/cm, and consequently, even if the dependence (3) could be observed, this would happen only in fields directly preceding breakdown.

Baraff's results<sup>[8]</sup> were obtained by numerically integrating the kinetic equation for a series of chosen values of the parameters and, in addition, pertained to temperatures that were sufficiently close to zero. The purpose of the present paper is to solve the problem of impact ionization in semiconductors in analytic form, and for arbitrary values of the field  $E$  and of the temperature  $T$ . We shall show below that in the region of large energies  $\epsilon \gg \hbar\omega$ , which are the only ones of interest for the impact-ionization problem, the distribution function of the electrons interacting with the acoustic and optical phonons in a covalent crystal is of the form

$$f_0(\epsilon) = \text{const} \cdot \epsilon^\nu \exp\left(-\frac{\epsilon}{eFl} s_0(E, T)\right), \quad (4)$$

where  $\epsilon$  is the electron energy, the parameter  $s_0(E, T)$  is determined as the positive root of the transcendental equation

$$(1 + \lambda) \text{ch} \frac{\hbar\omega}{2kT} \left/ \left[ \lambda \text{ch} \frac{\hbar\omega}{2kT} + \text{ch} \left( \frac{\hbar\omega}{2kT} - s_0 \frac{\hbar\omega}{eFl} \right) \right] \right. \\ \left. + \frac{1}{2s_0} \ln \frac{1 - s_0}{1 + s_0} = 0, \quad s_0 > 0, \quad (5)^*$$

and the exponent  $\nu$  is expressed in terms of  $s_0$ :

$$\nu = -\frac{1}{2} \left\{ 1 - \frac{\hbar\omega}{eFl} \text{th} \left( \frac{\hbar\omega}{2kT} \right) \frac{s_0}{1 - s_0^2} \right. \\ \times \left[ \frac{1 + \lambda}{1 - s_0^2} + \frac{1 + \lambda}{2s_0} \ln \frac{1 - s_0}{1 + s_0} \right. \\ \left. - s_0 \frac{\hbar\omega}{eFl} \left( \frac{1}{2s_0} \ln \frac{1 - s_0}{1 + s_0} \right)^2 \right. \\ \left. \times \text{sh} \left( \frac{\hbar\omega}{2kT} - s_0 \frac{\hbar\omega}{eFl} \right) \left/ \text{ch} \frac{\hbar\omega}{2kT} \right]^{-1} \right\}. \quad (6)^\dagger$$

From these formulas  $F$  is the effective electric field intensity, which coincides with its true value  $E$  in the case of a semiconductor with a scalar

effective carrier mass (independent of the direction). In the presence of anisotropy

$$F = (m/m_{\parallel})^{1/2} E, \quad (7)$$

where  $m$  is the effective mass, determined from the state density,

$$m = (m_x m_y m_z)^{1/3},$$

and  $m_{\parallel}$  is the effective mass for motion along the field direction

$$\frac{1}{m_{\parallel}} = \frac{\cos^2 \gamma_x}{m_x} + \frac{\cos^2 \gamma_y}{m_y} + \frac{\cos^2 \gamma_z}{m_z};$$

$m_x^{-1}$ ,  $m_y^{-1}$ , and  $m_z^{-1}$  are the principal values of the effective-mass tensor, and  $\gamma_x$ ,  $\gamma_y$ , and  $\gamma_z$  are the angles between the field  $E$  and the principal axes of this tensor. Thus, formula (7) contains, in particular, the anisotropy of the impact ionization and of the breakdown fields. The dimensionless parameter  $\lambda$  determines the relative contribution of the scattering by acoustic and optical phonons to the total mean free path

$$\lambda = \frac{l_{op}}{l_{ac}}, \quad \frac{1}{l} = \frac{1}{l_{op}} + \frac{1}{l_{ac}} = \frac{1 + \lambda}{l_{op}}.$$

We now consider qualitatively the behavior of the parameter  $s_0$  as the field intensity is varied. In extremely strong fields,  $eFl \gg \hbar\omega$ , we have

$$s_0 \approx \frac{3}{1 + \lambda} \frac{\hbar\omega}{eFl} \text{th} \frac{\hbar\omega}{2kT}.$$

In this case the distribution function (4) is close to that investigated by Wolff,<sup>[5]</sup> and consequently the diffusion approximation describes the real situation quite well. We assume now that  $2kT \ll \hbar\omega$ , and investigate the region of fields such that  $kT \lesssim eFl \ll \hbar\omega$ . In this region  $s_0$  is close to unity, but it never reaches this value, remaining smaller:

$$s_0 \approx 1 - 2 \exp \left\{ -2(1 + \lambda) \text{ch} \frac{\hbar\omega}{2kT} \left[ \lambda \text{ch} \frac{\hbar\omega}{2kT} \right. \right. \\ \left. \left. + \text{ch} \left( \frac{\hbar\omega}{2kT} - \frac{\hbar\omega}{eFl} \right) \right] \right\}.$$

When  $eFl \approx 2kT$  the value of  $s_0$  reaches its maximum

$$s_{0m} \approx 1 - 2 \exp \left\{ -\frac{2(1 + \lambda)}{\lambda + \text{sech}(\hbar\omega/2kT)} \right\}, \quad (8)$$

and in still smaller fields it begins to decrease. In this entire range of fields, when  $eFl \ll \hbar\omega$ , the distribution function (4) is proportional to  $\exp(-\epsilon/eFl)$ , i.e., it describes electrons accelerated by the field to an energy  $\epsilon$  without colli-

\*ch = cosh.

†th = tanh; sh = sinh.

sions, in other words, it corresponds to the picture considered by Shockley.<sup>[1]</sup> At the smallest fields  $eFl \ll kT$  the parameter  $s_0$  tends to  $eFl/kT$ , and the distribution function goes over into a Maxwellian distribution with a temperature equal to the lattice temperature.

If the effective-mass approximation, i.e., the assumption that the electron energy is quadratically dependent on the momentum, cannot be valid at energies on the order of  $\epsilon_i$ , then the mean free path can no longer be regarded as energy independent, i.e.,  $l = l(\epsilon)$ . In this case, as will be shown below, the distribution function  $f_0(\epsilon)$  becomes not Maxwellian, but of the form

$$f_0(\epsilon) \sim \exp \left\{ - \int_{\epsilon_0}^{\epsilon} \frac{d\epsilon'}{eFl(\epsilon')} s_0(\epsilon') \right\}, \quad (4a)$$

where  $\epsilon_0$  is the bottom of the band and  $s_0(\epsilon)$  is determined by the same transcendental equation (5), the only difference being that the constant  $l$  in it should be replaced by the function  $l(\epsilon)$ . The qualitative analysis presented above of the dependence  $s_0$  on the field intensity  $\mathbf{E}$  is consequently applicable also to this case, but for a quantitative comparison of the theory with experiment the difference between (4) and (4a) may turn out to be significant.

We present also a formula for the impact-ionization coefficient  $\kappa(\mathbf{E}, T)$ —the average number of ionization collisions of the electron per unit length of path. In the case described by (4) we have

$$\begin{aligned} \kappa(\mathbf{E}, T) = & a \left( \frac{m_{\parallel}}{m} \right)^{1/2} \frac{eF}{\epsilon_i} \left( \frac{\epsilon_i s_0}{eFl + \hbar\omega e^{(s_0-1)/s_0}} \right)^{\nu+2} \\ & \times \varphi_{\hbar} \left\{ \frac{\epsilon_i s_0}{eFl} \left( \frac{s_0^2}{12p} \right)^{1/\hbar} \right\} \exp \left( - \frac{\epsilon_i}{eFl} s_0 \right), \end{aligned} \quad (9)$$

where  $a$  is a numerical coefficient of the order of unity,  $\varphi_{\hbar}(z)$  a function defined by formula (75) below, and  $p$  and  $k$  are constants characterizing the energy dependence of the impact-ionization cross section near threshold, in accordance with formula (72). The asymptotic behavior of  $\varphi_{\hbar}(z)$  at both small and large values of the argument has a power-law character:

$$\begin{aligned} \varphi_{\hbar}(z) & \approx k! z^{-k}, & z \gg 1, \\ \varphi_{\hbar}(z) & \approx 2\Gamma \left( 2 \frac{k+1}{k+2} \right) \left( \frac{z}{k+2} \right)^{-k/(k+2)}, & z \ll 1. \end{aligned}$$

The following explanation is in order with respect to formula (9). The most important factors which are contained in it—the exponential function and  $\varphi_{\hbar}$ —follow from formulas (4) and (4a), and therefore are not subject to any doubt,

in contradistinction from the factor preceding them, which is essentially connected with the normalization constant in  $f_0(\epsilon)$ . The latter cannot be determined accurately when  $eFl \ll \hbar\omega$ , for then the greater part of the electrons have energies  $\epsilon \lesssim \hbar\omega$ , for which our solutions (4) and (4a) are generally inapplicable. Therefore the factor in the parentheses in (9) was obtained by interpolation between the limiting cases  $eFl \ll kT$ ,  $kT \ll eFl \ll \hbar\omega$ , and  $\hbar\omega \ll eFl$ . The specific form of the interpolating formula is of no importance, since the role of this entire term in (9) is very small compared with the two factors following it.

We now proceed to derive formulas (4)–(7). The electron momentum distribution function  $f(\mathbf{p})$  is determined by the usual kinetic equation

$$eV_{\mathbf{p}} f(\mathbf{p}) + S_{\mathbf{p}}^{-} f(\mathbf{p}) = S_{\mathbf{p}}^{+} \{f\}, \quad (10)$$

where  $\mathbf{p}$  is the electron momentum,  $S_{\mathbf{p}}^{-}$  the probability of collision with the phonon:

$$\begin{aligned} S_{\mathbf{p}}^{-} = & \frac{2\pi}{\hbar} \int |m_{\mathbf{k}}|^2 \{ (1 + N_{\omega_{\mathbf{k}}}) \delta(\epsilon_{\mathbf{p}} - \epsilon_{\mathbf{p}+\mathbf{k}} - \hbar\omega_{\mathbf{k}}) \\ & + N_{\omega_{\mathbf{k}}} \delta(\epsilon_{\mathbf{p}} - \epsilon_{\mathbf{p}+\mathbf{k}} + \hbar\omega_{\mathbf{k}}) \} \frac{d^3 k}{(2\pi\hbar)^3}, \end{aligned} \quad (11)$$

$m_{\mathbf{k}}$ —the matrix element of the interaction between the electron and the phonon with wave vector  $\mathbf{k}$ ,  $\omega_{\mathbf{k}}$  the phonon frequency, and  $N_{\omega_{\mathbf{k}}}$  the equilibrium number of such phonons:

$$N_{\omega_{\mathbf{k}}} = \left( \exp \frac{\hbar\omega_{\mathbf{k}}}{kT} - 1 \right)^{-1}, \quad (12)$$

$\epsilon_{\mathbf{p}}$  is the energy of an electron with momentum  $\mathbf{p}$ ;  $S_{\mathbf{p}}^{+} \{f\}$ —the number of electrons arriving per unit time in the state with momentum  $\mathbf{p}$  from all other states as a result of emission or absorption of phonons:

$$\begin{aligned} S_{\mathbf{p}}^{+} \{f\} = & \frac{2\pi}{\hbar} \int |m_{\mathbf{k}}|^2 \{ (1 + N_{\omega_{\mathbf{k}}}) \delta(\epsilon_{\mathbf{p}} - \epsilon_{\mathbf{p}+\mathbf{k}} + \hbar\omega_{\mathbf{k}}) \\ & + N_{\omega_{\mathbf{k}}} \delta(\epsilon_{\mathbf{p}} - \epsilon_{\mathbf{p}+\mathbf{k}} - \hbar\omega_{\mathbf{k}}) \} f(\mathbf{p} + \mathbf{k}) \frac{d^3 k}{(2\pi\hbar)^3}. \end{aligned} \quad (13)$$

We shall assume that there are two branches of phonons—acoustic and optical. The energy losses are essentially connected with the emission of optical phonons, for which we assume that

$$m_{\mathbf{k}} = m_0 = \text{const}, \quad \omega_{\mathbf{k}} = \omega = \text{const}.$$

The energy lost to emission of acoustical phonons can be neglected in the energy region of interest to us  $\epsilon_{\mathbf{p}} \gg \hbar\omega$ , i.e., we can leave out the  $\hbar\omega_{\mathbf{k}}$  of the acoustical phonons in the arguments of the  $\delta$  functions of (11) and (13). Using also the approximations customarily made for acoustical phonons

$$N_{\omega_{\mathbf{k}}} \approx 1 + N_{\omega_{\mathbf{k}}} \approx kT/\hbar\omega_{\mathbf{k}} \sim |\mathbf{k}|^{-1}, \quad |m_{\mathbf{k}}|^2 \sim |\mathbf{k}|,$$

we verify that only  $f(\mathbf{p} + \mathbf{k})$  and  $\epsilon(\mathbf{p} + \mathbf{k})$  in the integrand of (13) depend on  $\mathbf{k}$ . Going over therefore to a new integration variable  $\mathbf{p}' = \mathbf{p} + \mathbf{k}$  and integrating over the surface  $\epsilon_{\mathbf{p}'} = \text{const}$ , we obtain

$$S_{\mathbf{p}}^- = \frac{1}{\tau_{ac}(\mathbf{p})} + \frac{1}{\tau_{op}(\mathbf{p})} \left\{ \frac{e^\beta}{2 \text{ch} \beta} \frac{\Omega(\epsilon_{\mathbf{p}} - \hbar\omega)}{\Omega(\epsilon_{\mathbf{p}})} + \frac{e^{-\beta}}{2 \text{ch} \beta} \frac{\Omega(\epsilon_{\mathbf{p}} + \hbar\omega)}{\Omega(\epsilon_{\mathbf{p}})} \right\}, \quad (14)$$

$$S_{\mathbf{p}}^+ \{f\} = \frac{f_0(\epsilon_{\mathbf{p}})}{\tau_{ac}(\mathbf{p})} + \frac{1}{\tau_{op}(\mathbf{p})} \left\{ \frac{e^\beta \Omega(\epsilon_{\mathbf{p}} + \hbar\omega)}{2\Omega(\epsilon_{\mathbf{p}}) \text{ch} \beta} f_0(\epsilon_{\mathbf{p}} + \hbar\omega) + \frac{e^{-\beta} \Omega(\epsilon_{\mathbf{p}} - \hbar\omega)}{2\Omega(\epsilon_{\mathbf{p}}) \text{ch} \beta} f_0(\epsilon_{\mathbf{p}} - \hbar\omega) \right\}. \quad (15)$$

We have introduced here the following notation:  $\beta = \hbar\omega/2kT$ ;  $f_0(\epsilon_{\mathbf{p}})$ —symmetrical part of the distribution function:

$$f_0(\epsilon) = \frac{1}{\Omega(\epsilon)} \int f(\mathbf{p}) \delta(\epsilon - \epsilon_{\mathbf{p}}) \frac{d^3 p}{(2\pi\hbar)^3}, \quad (16)$$

$$\Omega(\epsilon) = \int \delta(\epsilon - \epsilon_{\mathbf{p}}) \frac{d^3 p}{(2\pi\hbar)^3}; \quad (17)$$

$\tau_{op}^{-1}(\mathbf{p})$ —frequency of collisions with the optical phonons:

$$\tau_{op}^{-1}(\mathbf{p}) = \frac{2\pi}{\hbar} |m_0|^2 \frac{\Omega(\epsilon_{\mathbf{p}})}{\text{th} \beta}; \quad (18)$$

$\tau_{ac}^{-1}(\mathbf{p})$ —frequency of collisions with acoustical phonons, determined in analogous fashion.

We now note that the energy dependence of  $\tau_{ac}^{-1}$  and  $\tau_{op}^{-1}$  is determined by the same factor  $\Omega(\epsilon)$ , and therefore their ratio does not depend on the energy:

$$\lambda = \frac{\tau_{op}(\mathbf{p})}{\tau_{ac}(\mathbf{p})} = \text{const}. \quad (19)$$

Using this circumstance and expanding (14) and (15) in powers of  $\hbar\omega/\epsilon_{\mathbf{p}}$  up to first order, we obtain

$$S_{\mathbf{p}}^- \approx \frac{1}{\tau(\mathbf{p})} \left\{ 1 - \frac{\hbar\omega}{1+\lambda} \frac{d \ln \Omega(\epsilon_{\mathbf{p}})}{d\epsilon_{\mathbf{p}}} \text{th} \beta \right\}, \quad (20)$$

$$S_{\mathbf{p}}^+ \{f\} \approx \frac{1}{(1+\lambda)\tau(\mathbf{p})} \left\{ \lambda + \frac{\text{ch}(\beta + \hbar\omega/d\epsilon_{\mathbf{p}})}{\text{ch} \beta} + \hbar\omega \frac{d \ln \Omega(\epsilon_{\mathbf{p}})}{d\epsilon_{\mathbf{p}}} \frac{\text{sh}(\beta + \hbar\omega/d\epsilon_{\mathbf{p}})}{\text{ch} \beta} \right\} f_0(\epsilon_{\mathbf{p}}). \quad (21)$$

In the last formula we have introduced, for convenience in notation, the operators  $\cosh(\beta + \hbar\omega/d\epsilon_{\mathbf{p}})$  and  $\sinh(\beta + \hbar\omega/d\epsilon_{\mathbf{p}})$ , whose action is defined by the well-known relations

$$\exp(\pm \hbar\omega/d\epsilon_{\mathbf{p}}) f_0(\epsilon_{\mathbf{p}}) \equiv f_0(\epsilon_{\mathbf{p}} \pm \hbar\omega).$$

The mean free time  $\tau(\mathbf{p})$  is connected with  $\tau_{ac}$  and  $\tau_{op}$  in the following fashion:

$$\tau^{-1}(\mathbf{p}) = \tau_{ac}^{-1}(\mathbf{p}) + \tau_{op}^{-1}(\mathbf{p}) = (1+\lambda)\tau_{op}^{-1}(\mathbf{p}). \quad (22)$$

We now solve (10) with respect to the distribution function  $f(\mathbf{p})$ , regarding  $S_{\mathbf{p}}^+ \{f\}$  as the free term of the equation

$$f(\mathbf{p}) = \int_{-\infty}^0 S_{\mathbf{p}+e\mathbf{E}t}^+ \{f\} \exp \left\{ \int_0^t S_{\mathbf{p}+e\mathbf{E}t'}^- dt' \right\} dt. \quad (23)$$

Substituting (20) and (21) in (23) and integrating the first two terms in  $S_{\mathbf{p}+e\mathbf{E}t}^+$  by parts, we obtain

$$\begin{aligned} f(\mathbf{p}) = & \left[ \lambda + \frac{\text{ch}(\beta + \hbar\omega/d\epsilon_{\mathbf{p}})}{\text{ch} \beta} \right] \frac{f_0(\epsilon_{\mathbf{p}})}{1+\lambda} \\ & - \int_{-\infty}^0 \exp \left\{ \int_0^t \left[ 1 - \hbar\omega \frac{d \ln \Omega(\epsilon_{\mathbf{p}+e\mathbf{E}t'})}{d\epsilon_{\mathbf{p}+e\mathbf{E}t'}} \frac{\text{th} \beta}{1+\lambda} \right] dt' \right\} \\ & \times \frac{dt'}{\tau(\mathbf{p}+e\mathbf{E}t')} \left\{ \frac{d\epsilon_{\mathbf{p}+e\mathbf{E}t}}{dt} \left[ \lambda + \frac{\text{ch}(\beta + \hbar\omega/d\epsilon_{\mathbf{p}+e\mathbf{E}t})}{\text{ch} \beta} \right] \right. \\ & \times \frac{f_0'(\epsilon_{\mathbf{p}+e\mathbf{E}t})}{1+\lambda} - \frac{\hbar\omega}{\tau(\mathbf{p}+e\mathbf{E}t)} \frac{d \ln \Omega(\epsilon_{\mathbf{p}+e\mathbf{E}t})}{d\epsilon_{\mathbf{p}+e\mathbf{E}t}} \\ & \times \left[ \frac{\text{th} \beta}{1+\lambda} \left( \lambda + \frac{\text{ch}(\beta + \hbar\omega/d\epsilon_{\mathbf{p}+e\mathbf{E}t})}{\text{ch} \beta} \right) \right. \\ & \left. \left. + \frac{\text{sh}(\beta + \hbar\omega/d\epsilon_{\mathbf{p}+e\mathbf{E}t})}{\text{ch} \beta} \right] \frac{f_0(\epsilon_{\mathbf{p}+e\mathbf{E}t})}{1+\lambda} \right\} dt. \quad (24) \end{aligned}$$

Here  $f_0'$  is the derivative of the function  $f_0$ .

Formula (24) is a generalization of the usual relation

$$f(\mathbf{p}) = f_0(\epsilon_{\mathbf{p}}) - e\mathbf{E}\tau(\mathbf{p}) \frac{d\epsilon_{\mathbf{p}}}{d\mathbf{p}} \frac{df_0(\epsilon_{\mathbf{p}})}{d\epsilon_{\mathbf{p}}}, \quad (25)$$

into which it goes over if we put  $e\mathbf{E}t \sim e\mathbf{E}\tau_{\mathbf{p}} \ll \mathbf{p}$  and  $\hbar\omega/\epsilon_{\mathbf{p}} \rightarrow 0$ . In our problem, however, in accordance with the statements made above, an important role may be played by electrons whose range accidentally is anomalously large, i.e.,  $t \gg \tau_{\mathbf{p}}$ . The velocity distribution of these electrons is certainly strongly anisotropic, and the approximation (25) is not suitable for their description, since it is based on the assumption that the asymmetrical part of the distribution function is much smaller than the symmetrical part.

Formula (24) reduces the problem of determining the total distribution function  $f(\mathbf{p})$  to the finding of its symmetrical part  $f_0(\epsilon_{\mathbf{p}})$ . In order to obtain an equation for  $f_0$ , we should average Eq. (24) over the constant-energy surface in accordance with the definition (16).

We present the subsequent calculations,

assuming the electron dispersion equation to be anisotropic but quadratic:

$$\epsilon_p = \frac{p_x^2}{2m_x} + \frac{p_y^2}{2m_y} + \frac{p_z^2}{2m_z}. \quad (26)$$

In this case our problem can be reduced to the problem with quadratic and isotropic dispersion

$$\epsilon_q = q^2/2m \quad (27)$$

by changing the scales along the  $x$ ,  $y$ , and  $z$  axes:

$$q_i = (m/m_i)^{1/2} p_i \quad (28)$$

( $i = x, y, z$ ) and introducing the effective electric field  $\mathbf{F}$ , whose components are connected with the components of the true field  $\mathbf{E}$  by relations analogous to (28):

$$F_i = (m/m_i)^{1/2} E_i, \quad (29)$$

$$F = (m/m_{||}) E. \quad (30)$$

For the average mass  $m$  we choose the effective mass determined from the energy-state density

$$m = (m_x m_y m_z)^{1/3}. \quad (31)$$

We now introduce for convenience in subsequent manipulations a set of new symbols. The effective mean free path  $l$  we define by the relation

$$l = \left(\frac{2\epsilon_p}{m}\right)^{1/2} \tau_p = \left(\frac{2\epsilon_q}{m}\right)^{1/2} \tau_q. \quad (32)$$

It is obvious that it is independent of the energy. All the energies are measured in units of  $eFl$ , i.e.,

$$\zeta_q = \frac{q^2}{2meFl}. \quad (33)$$

The optical-phonon energy expressed in these units will be denoted by  $\alpha$ :

$$\alpha = \hbar\omega/eFl. \quad (34)$$

Finally,  $u$  and  $v$  denote the cosines of the angles between the field  $\mathbf{F}$  and the vectors  $\mathbf{q}$  and  $\mathbf{q} + e\mathbf{F}t$  respectively:

$$u = \cos \{\mathbf{F}, \mathbf{q}\}, \quad v = \cos \{\mathbf{F}, \mathbf{q} + e\mathbf{F}t\}. \quad (35)$$

The quantities  $\zeta = \zeta_{\mathbf{q}}$ ,  $\zeta' = \zeta_{\mathbf{q}+e\mathbf{F}t}$ ,  $u$ , and  $v$  are connected by the obvious relation

$$\zeta(1-u^2) = \zeta'(1-v^2). \quad (36)$$

It follows from (17) and (33) that

$$\frac{d \ln \Omega(\zeta)}{d\zeta} = \frac{1}{2\zeta} \quad (37)$$

We now multiply (24) by  $\Omega^{-1}(\epsilon) \delta(\epsilon - \epsilon_p)$  and integrate it over all the momenta  $\mathbf{p}$ . Using then

formulas (26)–(37), we obtain the following equation for the function  $f_0(\zeta)$ :

$$\begin{aligned} & \left\{ 1 - \frac{1}{1+\lambda} \left[ \lambda + \frac{\text{ch}(\alpha + \beta d/d\zeta)}{\text{ch} \beta} \right] \right\} f_0(\zeta) + \frac{1}{2} \int_{-1}^1 du \int_{-1}^1 \frac{dv}{1+\lambda} \\ & \times \exp \left\{ - \int_v^u \left[ \frac{2\zeta(1-u^2)}{(1-\eta^2)^2} - \frac{\alpha \text{th} \beta}{(1+\lambda)(1-\eta^2)} \right] d\eta \right\} \\ & \times \left\{ \frac{2\zeta(1-u^2)}{(1-v^2)^2} v \left[ \lambda + \frac{\text{ch}(\beta + \alpha d/d\zeta')}{\text{ch} \beta} \right] f_0' \right. \\ & \times \left( \zeta \frac{1-u^2}{1-v^2} \right) - \frac{\alpha}{1-v^2} \left[ \frac{\text{th} \beta}{1+\lambda} \left( \lambda + \frac{\text{ch}(\beta + \alpha d/d\zeta')}{\text{ch} \beta} \right) \right. \\ & \left. \left. + \frac{\text{sh}(\beta + \alpha d/d\zeta')}{\text{ch} \beta} \right] f_0 \left( \zeta \frac{1-u^2}{1-v^2} \right) \right\} = 0. \quad (38) \end{aligned}$$

The derivative with respect to  $\zeta'$  in (38) denotes differentiation with respect to the total argument  $\zeta' = \zeta(1+u^2)/(1-v^2)$ .

Equation (38) was obtained by expanding with respect to  $\hbar\omega/\epsilon$ , and is consequently valid only in the region  $\epsilon \gg \hbar\omega$ , i.e.,  $\zeta \gg \alpha$ , and consequently it is meaningful to seek its solution only in this region. This is most conveniently done by taking the Laplace transform, i.e., going over to the function

$$\tilde{f}(s) = \int_0^\infty e^{-s\zeta} f_0(\zeta) d\zeta. \quad (39)$$

Then

$$f_0(\zeta) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{s\zeta} \tilde{f}(s) ds, \quad (40)$$

where  $c > \text{Re } s_0$ , and  $s_0$  is the singular point of  $\tilde{f}(s)$  farthest to the right. At large  $\zeta$ , the asymptotic value of  $f_0(\zeta)$  is determined only by the position and the character of the extreme right singular point  $s_0$ . Therefore, going over from (38) to the expression for  $\tilde{f}(s)$ , we should investigate only the singular points of this equation.

Here, however, it will be more convenient for us to write the equation not for  $\tilde{f}(s)$ , but for the related function  $\Phi(s)$ :

$$\begin{aligned} \Phi(s) &= \int_0^\infty \left\{ \lambda + \frac{\text{ch}(\beta + \alpha d/d\zeta)}{\text{ch} \beta} \right\} \frac{f_0(\zeta)}{1+\lambda} e^{-s\zeta} d\zeta \\ &= \left( \lambda + \frac{\text{ch}(\beta + \alpha s)}{\text{ch} \beta} \right) \frac{\tilde{f}(s)}{1+\lambda} \\ &\quad - \frac{e^{\beta+\alpha s}}{2(1+\lambda)\text{ch} \beta} \int_0^\infty e^{-s\zeta} f_0(\zeta) d\zeta. \quad (41) \end{aligned}$$

We multiply (38) by  $e^{-s\zeta}$  and integrate with respect to  $\zeta$  from zero to infinity, representing

$f_0(\xi')$  in the integrand in terms of  $\tilde{f}(s)$  with the aid of (40). After several elementary transformations we obtain for  $\Phi(s)$  the following equation:

$$\left[ \frac{(1 + \lambda) \operatorname{ch} \beta}{\lambda \operatorname{ch} \beta + \operatorname{ch}(\beta + \alpha s)} - \frac{1}{2s} \ln \frac{1 + s}{1 - s} \right] \Phi(s) - \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} [K(s, s')\Phi(s') - g(s, s')] ds' = h(s), \quad (42)$$

where

$$\begin{aligned} K(s, s') &= \frac{1}{2} \int_{-1}^1 \frac{du}{1-u^2} \int_{-1}^u \frac{dv}{1-v^2} \left( \frac{1+u}{1-u} \frac{1-v}{1+v} \right)^{\alpha \operatorname{th} \beta / 2(1+\lambda)} \\ &\times \left( \frac{s+u}{1-u^2} - \frac{s'+v}{1-v^2} + \frac{1}{2} \ln \frac{1+u}{1-u} \frac{1-v}{1+v} \right)^{-1} \\ &\times \left[ \frac{s'(1-v^2)}{(1+s'v)^2} + \frac{\alpha \operatorname{th} \beta}{(1+\lambda)(1+s'v)} \right. \\ &\left. + \frac{\alpha \operatorname{sh}(\beta + \alpha s')}{\lambda \operatorname{ch} \beta + \operatorname{ch}(\beta + \alpha s')} \right], \\ g(s, s') &= \frac{\alpha}{2} \int_{-1}^1 \frac{du}{1-u^2} \int_{-1}^u \frac{dv}{1-v^2} \left( \frac{1+u}{1-u} \frac{1-v}{1+v} \right)^{\alpha \operatorname{th} \beta / 2(1+\lambda)} \\ &\times \left( \frac{s+u}{1-u^2} - \frac{s'+v}{1-v^2} + \frac{1}{2} \ln \frac{1+u}{1-u} \frac{1-v}{1+v} \right)^{-1} \\ &\times \frac{e^{\beta + \alpha s'}}{2(1+\lambda) \operatorname{ch} \beta} \left[ 1 - \frac{\operatorname{sh}(\beta + \alpha s')}{\lambda \operatorname{ch} \beta + \operatorname{ch}(\beta + \alpha s')} \right] \\ &\times \int_0^\alpha f_0(\xi) e^{-s\xi} d\xi, \end{aligned} \quad (44)$$

$$\begin{aligned} h(s) &= \frac{1}{2} \int_1^\alpha \frac{\Phi(s') ds'}{s'(s-s')} - \frac{e^{\beta + \alpha s}}{2[\lambda \operatorname{ch} \beta + \operatorname{ch}(\beta + \alpha s)]} \\ &\times \int_0^\alpha f_0(\xi) e^{-s\xi} d\xi. \end{aligned} \quad (45)$$

In the derivation of (42) we assume that  $\operatorname{Re} s_0 < c < \operatorname{Re} s$ . We now note that the kernels  $K(s, s')$  and  $g(s, s')$  have as functions of  $s'$  a logarithmic cut along the real semi-axis  $(s, +\infty)$  in the half-plane  $\operatorname{Re} s' > c$ , and also an additional branch point  $s' = 1$ . The points determined by the zeros of the expression

$$\lambda \operatorname{ch} \beta + \operatorname{ch}(\beta + \alpha s),$$

are not singular for  $\Phi(s)$ , as can be readily verified by multiplying all the equations in (42) by this expression. We then see directly that at these points

$$(1 + \lambda) \operatorname{ch} \beta \Phi(s) + \frac{1}{2} e^{\beta + \alpha s} \int_0^\alpha f_0(\xi) e^{-s\xi} d\xi = 0,$$

and consequently the singularities corresponding to them in  $K(s, s')\Phi(s')$  and  $g(s, s')$  cancel out.

We consider further the values of  $s$  in the strip  $-1 < \operatorname{Re} s < 1$ , since  $s = 1$  is likewise not a singular point of  $\Phi(s)$ . The singularities of the different terms of the equation at these points cancel each other. Thus, for example, the logarithm in the first term in the left side of (42) cancels a similar divergence of the integral in the first term of (45). Shifting the contour of integration in (42) to the right, we reduce this equation to the form

$$\left[ \frac{(1 + \lambda) \operatorname{ch} \beta}{\lambda \operatorname{ch} \beta + \operatorname{ch}(\beta + \alpha s)} - \frac{1}{2s} \ln \frac{1 + s}{1 - s} \right] \Phi(s) - \int_s^\infty \kappa(s, s') \Phi(s') ds' = h(s) - \int_s^\infty \gamma(s, s') ds', \quad (46)$$

where  $\kappa(s, s')$  and  $\gamma(s, s')$  are the jumps in the functions  $K(s, s')$  and  $g(s, s')$  on the cut  $s \leq s' < \infty$ .

On the segment  $s \leq s' < 1$  we have

$$\begin{aligned} \kappa(s, s') &= \frac{1}{2\pi i} [K(s, s' + i\delta) - K(s, s' - i\delta)] \\ &= \frac{1}{2} \int_{-1}^1 \frac{du}{1-u^2} \int_{-1}^u \frac{dv}{1-v^2} \left( \frac{1+u}{1-u} \frac{1-v}{1+v} \right)^{\alpha \operatorname{th} \beta / 2(1+\lambda)} \\ &\times \left[ \frac{s'(1-v^2)}{(1+s'v)^2} + \frac{\alpha \operatorname{th} \beta}{(1+\lambda)(1+s'v)} \right. \\ &\left. + \frac{\alpha \operatorname{sh}(\beta + \alpha s')}{\lambda \operatorname{ch} \beta + \operatorname{ch}(\beta + \alpha s')} \right] \\ &\times \delta \left( \frac{s+u}{1-u^2} - \frac{s'+v}{1-v^2} + \frac{1}{2} \ln \frac{1+u}{1-u} \frac{1-v}{1+v} \right). \end{aligned} \quad (47)$$

The terms in the left side of (46) have singularities at those points where the function  $\Phi(s)$  has singularities, while the terms in the right side have no singularities at these points.

The only singular point of  $\Phi(s)$  in the half-plane  $\operatorname{Re} s > -1$  is determined by the zero of the coefficient of  $\Phi(s)$  in the left side of (46):

$$\frac{(1 + \lambda) \operatorname{ch} \beta}{\lambda \operatorname{ch} \beta + \operatorname{ch}(\beta + \alpha s_0)} - \frac{1}{2s_0} \ln \frac{1 + s_0}{1 - s_0} = 0. \quad (48)$$

The point  $s_0$  is also the farthest singularity of  $\Phi(s)$  on the right. In order to investigate the behavior of  $\Phi(s)$  in the vicinity of  $s_0$ , we expand all the coefficients in (46) in powers of  $(s - s_0)$  near this point, retaining the first nonvanishing term. This equation then takes the form

$$(s - s_0) \Phi(s) + \nu \int_s^\infty \Phi(s') ds' = \text{const}, \quad (49)$$

where

$$v = \kappa(s_0, s_0) \left\{ \frac{d}{ds_0} \left[ \frac{(1+\lambda) \operatorname{ch} \beta}{\lambda \operatorname{ch} \beta + \operatorname{ch}(\beta + \alpha s_0)} - \frac{1}{2s_0} \ln \frac{1+s_0}{1-s_0} \right] \right\}^{-1}. \quad (50)$$

The solution of (49) is of the form  $\operatorname{const} \cdot (s - s_0)^{-\nu-1}$ . More strictly speaking this denotes that  $\Phi(s)$  is of the form

$$\Phi(s) = (s - s_0)^{-\nu-1} \varphi_1(s) + \varphi_2(s), \quad (51)$$

where  $\varphi_1(s)$  and  $\varphi_2(s)$  are functions that are regular at the point  $s_0$ . They cannot be obtained without a complete solution of Eq. (46). For our purposes, however, this is not necessary.

Equation (48), which determines  $s_0$ , has two roots. One is  $s_0 = 0$  and the other  $s_0 < 0$ . However, the singular point corresponds only to the second, negative root, since  $\nu = -1$  when  $s_0 = 0$ , and consequently the function (51) is regular at this point. Indeed,

$$\kappa(s, s) = \frac{1}{2s} \left[ \frac{1}{1-s^2} - \frac{1}{2s} \ln \frac{1+s}{1-s} \right] + \frac{\alpha \operatorname{th} \beta}{2(1+\lambda)} \frac{1}{1-s^2} + \frac{\alpha}{4s} \ln \frac{1+s}{1-s} \frac{\operatorname{sh}(\beta + \alpha s)}{\lambda \operatorname{ch} \beta + \operatorname{ch}(\beta + \alpha s)}. \quad (52)$$

Substituting this result in (50), we obtain

$$v = -\frac{1}{2} \left\{ 1 + \alpha \frac{s_0 \operatorname{th} \beta}{1-s_0^2} \left[ \frac{1+\lambda}{1-s_0^2} - \frac{1+\lambda}{2s_0} \ln \frac{1+s_0}{1-s_0} \right] + \alpha s_0 \frac{\operatorname{sh}(\beta + \alpha s_0)}{\operatorname{ch} \beta} \left( \frac{1}{2s_0} \ln \frac{1+s_0}{1-s_0} \right)^2 \right\}^{-1}. \quad (53)$$

The point  $s_0$  is also the extreme right singularity for the function  $\tilde{f}(s)$ . Substituting (51) and (41) in (40) and shifting the contour of integration in (40) to the left, we obtain the following asymptotic representation for  $f_0(\zeta)$  at large  $\zeta$ :

$$f_0(\zeta) = C \zeta^{\nu} e^{s_0 \zeta}. \quad (54)$$

The constant  $C$  is proportional to  $\varphi_1(s_0)$  and therefore cannot be calculated in the approximation employed by us. However, its influence on the probability of impact ionization is negligibly small, and we confine ourselves to an estimate of its order of magnitude. This is easily done on the basis of the following considerations. Physically,  $C$  is connected with the normalization constant in the distribution function. In the case of large fields  $\alpha \ll 1$ , when most electrons have an energy on the order of  $\alpha^{-1} eFl \gg eFl \gg \hbar\omega$  and are consequently in the region of applicability of the solution (54),  $C$  simply coincides with the normalization constant

$$C = n \frac{(2\pi^3)^{1/2} \hbar^3}{m^{3/2}} \left[ \frac{3}{1+\lambda} \frac{\hbar\omega}{e^2 F^2 l^2} \operatorname{th} \frac{\hbar\omega}{2kT} \right]^{3/2} \sim \left( \frac{\hbar}{m\omega} \frac{3\alpha^2}{1+\lambda} \operatorname{th} \beta \right)^{3/2} n, \quad (55)$$

where  $n$  is the total number of electrons.

If  $\alpha \gg 1$ , then the bulk of the electrons has energies smaller than  $\hbar\omega$  and is not described by the function (54). In this case  $C$  is the number of electrons in the tail of the distribution function, and can be estimated in the following fashion. In this case the electrons fill practically uniformly the region of energies  $\epsilon_{\mathbf{p}} < \hbar\omega$ , and  $n(\epsilon)$ , the number of electrons reaching an energy  $\epsilon$ , is proportional to the probability of passing without collision through the energy range from  $\epsilon_{\mathbf{p}} < \hbar\omega$  to  $\epsilon$ . This probability is in turn proportional to

$$\exp\left(-\int_0^t S_{\mathbf{p}+e\mathbf{E}t'} dt'\right),$$

where  $t$  is determined from the condition  $\epsilon_{\mathbf{p}+e\mathbf{E}t} = \epsilon$ . An elementary calculation of this type yields under the condition  $1 \ll \alpha \lesssim \beta$

$$n(\zeta) \sim n\alpha^{-1} \exp\left\{-\zeta + \frac{\alpha \operatorname{th} \beta}{2(1+\lambda)} \ln \zeta - \frac{\alpha \operatorname{th} \beta}{2(1+\lambda)} \left( \ln \frac{\alpha}{4} - 1 \right)\right\}. \quad (56)$$

On the other hand, the quantity  $n(\zeta)$  can be determined also as

$$n(\zeta) = \frac{\sqrt{2}(meFl)^{3/2}}{\pi^2 \hbar^3} \int_{\zeta}^{\infty} f_0(\xi) \sqrt{\xi} d\xi.$$

Substituting here (54) and using the fact that  $s_0 \rightarrow -1$  when  $\alpha \gg 1$ , and that the exponent

$$v \approx -\frac{1}{2} + \frac{\alpha}{2(1+\lambda)} \operatorname{th} \beta,$$

we obtain

$$n(\zeta) \approx \frac{\sqrt{2}}{\pi^2} \left( \frac{m\omega}{\hbar\alpha} \right)^{3/2} C \exp\left\{-\zeta + \frac{\alpha \operatorname{th} \beta}{2(1+\lambda)} \ln \zeta\right\}. \quad (57)$$

Comparing (56) and (57) we obtain, leaving out numerical factors,

$$C \sim n \left( \frac{\hbar}{m\omega} \right)^{3/2} \sqrt{\alpha} \exp\left\{-\frac{\alpha \operatorname{th} \beta}{2(1+\lambda)} \left( \ln \frac{\alpha}{4} - 1 \right)\right\}. \quad (58)$$

It is now easy to construct an interpolation formula for  $C$ , which when  $\alpha \ll 1$  coincides with (55), and when  $\alpha \gg 1$  with (58); in the intermediate region  $\alpha \sim 1$  the formula gives a result of the correct order of magnitude. For example,

$$C(\alpha, \beta) \sim -n \frac{s_0}{\alpha} \left( -\frac{\hbar}{m\omega} \alpha s_0 \right)^{3/2} \times \exp \left\{ -\frac{\alpha \operatorname{th} \beta}{2(1+\lambda)} \left( \ln \frac{\alpha}{4} - 1 \right) \right\}. \quad (59)$$

We emphasize once more that a more accurate calculation of  $C(\alpha, \beta)$  is of no interest in the problem of impact ionization, for in the worst case  $C(\alpha, \beta)$  results in corrections to the exponential in (54), the relative magnitude of which is of the order of  $\hbar\omega/\epsilon_1 \sim 10^{-2}$ .

More important are the other simplifications made in the derivation of (54), particularly the assumption that  $\epsilon_p$  is a quadratic function of the momentum up to an energy on the order of  $\epsilon_1$ , and that the influence of the impact ionization process itself on the function  $f_0(\epsilon)$  in the region  $\epsilon \leq \epsilon_1$  is small. The first of these limitations can be eliminated. We now present a derivation which is somewhat less rigorous than the preceding one, but which leads to the same results and makes it possible to take into account the non-parabolic nature of the function  $\epsilon_p$ . For concreteness we confine ourselves here to the most interesting case, when the dispersion law is not parabolic but isotropic:

$$\epsilon_p = \Delta \left( 1 + \frac{p^2}{m\Delta} \right)^{1/2}, \quad (60)$$

where  $\Delta$  is the half-width of the forbidden band, from the middle of which the energies are reckoned in the given case. This type of dependence of  $\epsilon_p$  is precisely the one which is typical, as is well known, for the majority of semiconductors of the  $A^{III}B^V$  group and for a number of other semiconductors. In this case

$$\Omega(\epsilon) = \frac{m^{3/2} \Delta^{1/2}}{2\pi^2 \hbar^3} \frac{\epsilon}{\Delta} \left( \frac{\epsilon^2}{\Delta^2} - 1 \right)^{1/2}$$

and the mean free path  $l(\epsilon)$  depends on  $\epsilon$  like

$$l(\epsilon) = \frac{\epsilon^2}{\Delta^2} l, \quad (61)$$

where  $l$  is the mean free path in the region of energies  $\hbar\omega \ll (\epsilon^2 - \Delta^2)^{1/2} \ll \Delta$ .

Since we are interested only in calculating the exponential in the distribution function  $f_0(\epsilon)$ , we now omit in (24) all the terms of order  $\hbar\omega/\epsilon$ . We change over again to the variables

$$\epsilon = \epsilon_p, \quad \epsilon' = \epsilon_p + eEt;$$

$$u = \cos \{ \mathbf{E}, \mathbf{p} \}, \quad v = \cos \{ \mathbf{E}, \mathbf{p} + eEt \}$$

$$= \left[ 1 - \frac{\epsilon^2 - \Delta^2}{\epsilon'^2 - \Delta^2} (1 - u^2) \right]^{1/2}. \quad (62)$$

After averaging (24) over the angles we obtain

$$f_0(\epsilon) = \left[ \lambda + \frac{\operatorname{ch}(\beta + \hbar\omega d/d\epsilon)}{\operatorname{ch} \beta} \right] \frac{f_0(\epsilon)}{1 + \lambda} - \frac{1}{2} \int_{-1}^1 du \int_{\Gamma}^{\epsilon} \exp \left\{ - \int_{\epsilon'}^{\epsilon} \left[ 1 - \frac{\epsilon^2 - \Delta^2}{\epsilon'^2 - \Delta^2} (1 - u^2) \right]^{-1/2} \right. \\ \left. \times \frac{d\epsilon''}{eEl(\epsilon'')} \right\} \left[ \lambda + \frac{\operatorname{ch}(\beta + \hbar\omega d/d\epsilon')}{\operatorname{ch} \beta} \right] \frac{f_0'(\epsilon')}{1 + \lambda} d\epsilon' \quad (63)$$

The integral with respect to  $\epsilon'$  in (62) is taken along the contour  $\Gamma$  which goes from  $+\infty$  along the real axis, but in the lower half-plane, to the branch point  $\epsilon_B'^2 - \Delta^2 = (\epsilon^2 - \Delta^2)(1 - u^2)$ , turning around this point and then proceeding to  $\epsilon' = \epsilon$ . The point  $\epsilon$  itself is assumed to lie on the upper branch of the contour  $\Gamma$ , if  $u > 0$  and on the lower if  $u < 0$ . When  $\epsilon'$  varies along the contour  $\Gamma$ , the quantity  $v$  (62) increases monotonically from  $-1$  to  $u$ . The branch point corresponds to  $v = 0$ . The integral with respect to  $\epsilon''$  is taken along a segment of the same contour  $\Gamma$ .

We now seek a solution of (63) in the form

$$f_0(\epsilon) = \text{const} \cdot \exp \left\{ - \int_{\Delta}^{\epsilon} s(\epsilon') \frac{d\epsilon'}{eEl(\epsilon')} \right\}. \quad (64)$$

Substituting (64) in (63), we shall assume that

$$\frac{d^n}{d\epsilon^n} f_0(\epsilon) \approx \left[ - \frac{s(\epsilon)}{eEl(\epsilon)} \right]^n f_0(\epsilon),$$

since the other terms which appear upon differentiation will be on the order of  $eEl/\epsilon$  or smaller, and we neglect terms of this order, because we seek only the asymptotic value of  $f_0(\epsilon)$  when  $\epsilon \gg eEl$ . As a result we obtain from (63) the following equation for  $s(\epsilon)$ :

$$\frac{\operatorname{ch} \beta - \operatorname{ch}[\beta - \alpha(\epsilon)s(\epsilon)]}{(1 + \lambda)\operatorname{ch} \beta} - \frac{1}{2} \int_{-1}^1 du \int_{\Gamma}^{\epsilon} \frac{d\epsilon'}{eEl(\epsilon')} \times \exp \left\{ - \int_{\epsilon'}^{\epsilon} \left[ 1 - \frac{\epsilon^2 - \Delta^2}{\epsilon'^2 - \Delta^2} (1 - u^2) \right]^{-1/2} \right. \\ \left. - s(\epsilon'') \right\} \frac{d\epsilon''}{eEl(\epsilon'')} \left[ \lambda + \frac{\operatorname{ch}[\beta - \alpha(\epsilon')s(\epsilon')]}{\operatorname{ch} \beta} \right] \frac{s(\epsilon')}{1 + \lambda} = 0. \quad (65)$$

The value of  $s(\epsilon)$ , as we shall show below, is always less than unity. Therefore

$$\frac{1}{v} - s(\epsilon'') = \left[ 1 - \frac{\epsilon^2 - \Delta^2}{\epsilon'^2 - \Delta^2} (1 - u^2) \right]^{-1/2} - s(\epsilon'') > 0$$

for  $v > 0$ , i.e., on the upper edge of the cut. If  $v < 0$ , i.e., on the lower edge of the cut, this difference is negative.

Thus, the exponential in (65) contains a negative quantity, the absolute value of which increases

monotonically with increasing distance between the points  $\epsilon'$  and  $\epsilon$ , and consequently the main contribution to the integral with respect to  $\epsilon'$  is made in the asymptotic region  $\epsilon \gg eEl$  by the region  $\epsilon' \sim \epsilon$ . Taking this circumstance into account, we expand the exponential in powers of  $(\epsilon - \epsilon')$  up to first order inclusive:

$$\int_{\epsilon'}^{\epsilon} \left[ \left( 1 - \frac{\epsilon^2 - \Delta^2}{\epsilon'^2 - \Delta^2} (1 - u^2) \right)^{-1/2} - s(\epsilon'') \right] \frac{d\epsilon''}{eEl(\epsilon'')} \\ \approx \left( \frac{1}{u} - s(\epsilon) \right) \frac{\epsilon - \epsilon'}{eEl(\epsilon)},$$

and replace the remaining quantities that depend on  $\epsilon'$  by their values at  $\epsilon' = \epsilon$ . We note furthermore that from the definition of the contour  $\Gamma$ , the difference  $\epsilon' - \epsilon$  is positive when  $u < 0$  and negative when  $u > 0$  in the region making the main contribution to the integral. Carrying out now all the integrations, we reduce (65) to the form

$$\frac{(1 + \lambda) \operatorname{ch} \beta}{\lambda \operatorname{ch} \beta + \operatorname{ch} [\beta - \alpha(\epsilon) s(\epsilon)]} - \frac{1}{2s(\epsilon)} \ln \frac{1 + s(\epsilon)}{1 - s(\epsilon)} = 0. \quad (66)$$

Thus, an account of the non-parabolic nature of the dispersion law introduces nothing essentially new to the problem of impact ionization, but makes the distribution function (64) non-Maxwellian.

It is easy to include in the scheme under consideration also the impact ionization process. To this end we must add to  $S_{\mathbf{p}}^{-1}$  (20) a term in the form  $\tau_{\mathbf{i}}^{-1}(\mathbf{p})$ , describing the ionization collision probability of an electron with momentum  $\mathbf{p}$ . The subsequent derivations are perfectly analogous to (63)–(66), so that we present immediately their final result. The distribution function can again be represented in the form (64), but with  $l(\epsilon)$ , the total mean free path, determined by the scattering by phonons and by the ionization collisions:

$$l(\epsilon) = \left( \frac{2\epsilon}{m} \right)^{1/2} \tau(\epsilon), \quad \tau^{-1}(\epsilon) = \tau_{ac}^{-1}(\epsilon) + \tau_{op}^{-1}(\epsilon) + \tau_{\mathbf{i}}^{-1}(\epsilon), \quad (67)$$

while the parameter  $s(\epsilon)$  is determined by an equation which differs somewhat from (66):

$$\frac{1 + \lambda}{1 - \mu(\epsilon)} \frac{\operatorname{ch} \beta}{\lambda \operatorname{ch} \beta + \operatorname{ch} [\beta - \alpha(\epsilon) s(\epsilon)]} - \frac{1}{2s(\epsilon)} \ln \frac{1 + s(\epsilon)}{1 - s(\epsilon)} = 0; \quad (68)$$

$\mu(\epsilon)$  is the relative contribution of the ionization collisions

$$\mu(\epsilon) = \tau(\epsilon) / \tau_{\mathbf{i}}(\epsilon). \quad (69)$$

At energies  $\epsilon < \epsilon_{\mathbf{i}}$  this ratio is equal to zero, and then (68) coincides with (66).

An experimentally determined characteristic of the impact ionization process is usually the coefficient of impact ionization  $\kappa(\mathbf{E}, T)$ , defined as the ratio of the average probability of impact ionization to the average electron drift velocity  $v_d$ :

$$\kappa(\mathbf{E}, T) = \frac{1}{nv_d} \int_0^{\infty} \tau_i^{-1}(\epsilon) f_0(\epsilon) \Omega(\epsilon) d\epsilon \equiv \frac{\sigma(\mathbf{E}, T)}{nv_d} f_0(\epsilon_{\mathbf{i}}). \quad (70)$$

For a distribution function in the form (64)  $\sigma(\mathbf{E}, T)$  is of the form

$$\sigma(\mathbf{E}, T) = \int_{\epsilon_{\mathbf{i}}}^{\infty} \tau_i^{-1}(\epsilon) \exp \left\{ - \int_{\epsilon_{\mathbf{i}}}^{\epsilon} s(\epsilon') \frac{d\epsilon'}{eEl(\epsilon')} \right\} \Omega(\epsilon) d\epsilon. \quad (71)$$

Compared with  $f_0(\epsilon_{\mathbf{i}})$ , this quantity, like  $v_d$ , is a slowly varying function of the field.

The main contribution to (71) is made by the region of energies close to threshold:  $\epsilon - \epsilon_{\mathbf{i}} \lesssim eEl(\epsilon_{\mathbf{i}}) s^{-1} \ll \epsilon_{\mathbf{i}}$ . In this region the probability of impact ionization can be represented in the form

$$\tau_i^{-1}(\epsilon) = \tau^{-1}(\epsilon_{\mathbf{i}}) p \left( \frac{\epsilon - \epsilon_{\mathbf{i}}}{\epsilon_{\mathbf{i}}} \right)^k, \quad (72)$$

where  $p$  is a dimensionless constant which, generally speaking, is much larger than unity,<sup>[9]</sup> and the exponent  $k$  can assume values equal to 1, 2, and 3 depending on whether the crystal is isotropic or not<sup>[10]</sup> and depending on how large its dielectric constant is.<sup>[9]</sup>

In order to calculate (71) in explicit form, we now note the following: in a narrow region of energies close to a threshold, which makes the main contribution to (71), the condition  $\mu(\epsilon) \ll 1$  is apparently always satisfied. Therefore, in the calculation of  $\sigma(\mathbf{E}, T)$ , we can put  $l(\epsilon) \approx l(\epsilon_{\mathbf{i}})$  and

$$s(\epsilon) = \frac{1}{2} [s(\epsilon_{\mathbf{i}}) + (s^2(\epsilon_{\mathbf{i}}) + 12\mu(\epsilon))^{1/2}]. \quad (73)$$

Indeed, expression (73) follows directly from (68), if

$$a(\epsilon) \approx a(\epsilon_{\mathbf{i}}) \ll 1$$

and  $\mu \ll 1$ . On the other hand, if

$$a(\epsilon) \approx a(\epsilon_{\mathbf{i}}) \gg \sqrt{3\mu(\epsilon)}$$

the quantity  $\mu$  in (68) plays no role at all and  $s(\epsilon) \approx s(\epsilon_{\mathbf{i}})$ . By virtue of the condition  $\mu \ll 1$ , these regions overlap and consequently (73) holds true everywhere. Using (72) and (73), we can easily transform (71) into

$$\sigma(\mathbf{E}, T) = \frac{\Omega(\epsilon_{\mathbf{i}})}{\tau(\epsilon_{\mathbf{i}})} eEl(\epsilon_{\mathbf{i}}) \varphi_h \left\{ \frac{\epsilon_{\mathbf{i}} s(\epsilon_{\mathbf{i}})}{eEl(\epsilon_{\mathbf{i}})} \left( \frac{s^2(\epsilon_{\mathbf{i}})}{12p} \right)^{1/h} \right\}, \quad (74)$$

$$\varphi_h(z) = z \int_0^{\infty} \exp \left\{ -\frac{z}{2} \int_0^x [1 + (1 + y^h)^{1/2}] dy \right\} x^h dx. \quad (75)$$

The calculation of the normalization constant in  $f_0(\epsilon)$  and of the drift velocity  $v_d$  in (70) is less definite. Both terms, however, vary very little as functions of  $E$ , compared with the exponential and with  $\varphi_k$ , so that it is sufficient to be able to write down at least their order of magnitude. This can be done by writing down their correct values in two limiting cases:  $eFl \ll kT$  and  $eFl \gg \hbar\omega$ , when the propagation function is almost isotropic and the function (4) gives a solution which describes correctly the majority of the electrons, after which one can use (59) to interpolate in the intermediate region  $kT \ll eFl \ll \hbar\omega$ . This method yielded the factor

$$\left( \frac{\epsilon_i s_0}{eFl + \hbar\omega e^{(s_0-1)/s_0}} \right)^{v+2}$$

in (9). As a rule, impact ionization is observed in such fields that  $\alpha \lesssim 1$ . Then, apart from factors of the order of unity, this term can be calculated from the formulas that pertain to the case  $\alpha \ll 1$ . In formula (9) this reduces simply to the fact that it is necessary to put  $\hbar\omega \exp[(s_0 - 1)/s_0] = 0$ .

Formulas (4)–(9) thus yield the solution of the problem of impact ionization in covalent semiconductors, for which as is well known,  $m_k = \text{const}$ , at arbitrary values of the field and of the temperature, and also under rather general assumptions concerning the character of the band structure and the interaction between the electrons and the phonons. There is, however, one limitation, which has been used in the derivation: we have assumed that the probability of scattering by an

optical phonon will not depend on the scattering angle, i.e.,  $m_k = \text{const}$ . This is precisely why we can write a closed equation such as (38) for the function  $f_0(\epsilon)$ , dependent only on the energy. Qualitative deductions concerning the variation of the law of the growth of the number of ionizing electrons with the field in fields such that  $eEl \sim \hbar\omega$  apparently remain also in force when  $m_k \neq \text{const}$ , but the quantitative estimates may differ. In particular, some caution is necessary when the results of the present work are applied to semiconductors with a noticeable fraction of ionic bonding, for in this case  $m_k$  contains a term proportional to  $|k|^{-1}$  for sufficiently small phonon momenta, i.e., small scattering angles.

<sup>1</sup>W. Shockley, Czech. J. Phys. **B11**, 81 (1961).

<sup>2</sup>A. G. Chynoweth, Phys. Rev. **109**, 1537 (1959).

<sup>3</sup>M. J. Druyvesteyn and F. M. Penning, Revs. Modern Phys. **12**, 87 (1940).

<sup>4</sup>B. I. Davydov, JETP **7**, 1069 (1937).

<sup>5</sup>P. A. Wolff, Phys. Rev. **95**, 1415 (1954).

<sup>6</sup>A. G. Chynoweth, J. Appl. Phys. **31**, 1161 (1960). Logan, Chynoweth, and Cohen, Phys. Rev. **128**, 2518 (1962).

<sup>7</sup>J. L. Moll and R. Vanoverstraeten, Electronics **6**, 147 (1963).

<sup>8</sup>G. A. Baraff, Phys. Rev. **128**, 2507 (1962).

<sup>9</sup>L. V. Keldysh, JETP **37**, 713 (1959), Soviet Phys. JETP **10**, 509 (1960).

<sup>10</sup>W. Franz, Encycl. Phys. **17**, 155 (1956).

Translated by J. G. Adashko