Theory of Electric Breakdown in the High-Temperature Region and Its Relation to Thermionic Emission

KURT LEHOVEC
Sprague Electric Company, North Adams, Massachusetts
(Received August 5, 1954)

The decrease of von Hippel's breakdown field caused by random fluctuations of electron energy is estimated from consideration of the distribution function of electrons in strong fields. This decrease of breakdown field is appreciable at ordinary temperatures in case of the interaction of electrons with optical modes of lattice vibrations in ionic lattices and this may account for the observed decrease of breakdown field with increasing temperatures at elevated temperatures.

The rate of formation of "fast" electrons from "slow" electrons is used to estimate the deviations from thermal equilibrium distribution of electrons at the surface of a thermionic cathode. Deviations from Richardson's equation are expected if $E(h\nu)/(kT)^2 \lesssim (3/16)^{\frac{1}{2}}$, where E is the height of energy barriers at the surface above the bottom of the conduction band in semiconductors, $h\nu$ is the phonon energy, and kT is the energy equivalent of temperature. A strong increase of the thermionic emission current with the field in the cathode is predicted for semiconductors at fields of the order of the breakdown field.

INTRODUCTION

ELECTRIC breakdown in insulators may arise by two different processes: (1) At high field-strength electrons from the filled band may enter the empty band by the wave mechanical tunnel effect (Zener effect).[1] (2) Electrons in the nearly empty band may acquire sufficient energy to release secondary electrons from the filled band by collision, these in turn release tertiary electrons, etc. (breakdown by electron avalanches).[2]

In this paper, we are concerned with the breakdown by electron avalanches only. The early theoretical work in this field[3] was based on considerations of the rate at which an average electron reaches sufficient energy to create a secondary electron. Seitz[4] has considered the effect of deviations from the average behavior of an electron on the breakdown field. More recently, Fröhlich,[5] Heller,[6] Franz,[7] and Yamashita and Watanabe[8] have discussed breakdown phenomena in terms of the distribution function of electrons over the energy states. In this work, the breakdown field is derived from the condition that the rate of creation of secondary electrons exceeds the rate of recombination of electrons excited by impact. Kawamura[9] has used an expression derived by Franz[7] for a comparison of the effect of random processes on breakdown field as obtained from the theory based on the distribution function with the expression derived by Seitz.[4]

In this paper, we would like to use considerations of the distribution function in strong fields for a discussion of (1) the relation of the breakdown theory based on the distribution function to the older one, based on the behavior of an individual electron; (2) the temperature dependence of the breakdown field at elevated temperatures; (3) a relation between breakdown field and thermionic emission. Instead of specializing the rather complicated expressions derived by Heller,[6] for our purposes it may be permissible to present a simplified derivation which involves a boundary condition of the distribution function at the ionization energy which differs from that of previous workers. For a justification of some of the applications made, the reader will be referred frequently to the paper by Heller.[6]

THE BEHAVIOR OF THE ELECTRONS AND THE EFFECT OF RANDOM PROCESSES

The release of a secondary electron by impact of a fast electron is pictured schematically in Fig. 1. A fast electron of energy E in the conduction band of the crystal excites an electron from the filled band to the conduction band and loses its own energy in this process. Thus by the impact of a fast electron, two slow electrons

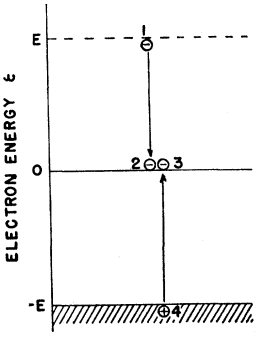

FIG. 1. Production of a secondary electron by impact. (1) Fast electron before impact. (2) Same electron after having lost its kinetic energy by impact. (3) Secondary electron excited by impact from the filled band into the conduction band. (4) Hole left behind in the filled band. $\epsilon > 0$: electron conduction band; $\epsilon < -E$: filled band.

[1] C. M. Zener, Proc. Roy. Soc. (London) **A145**, 523 (1934); W. Franz, Z. Physik **113**, 607 (1939); W. V. Houston, Phys. Rev. **57**, 184 (1940).
[2] See A. von Hippel, Trans. Faraday Soc. **A42**, 78 (1946), and A. von Hippel and R. S. Alger, Phys. Rev. **76**, 127 (1949) for recent accounts of this point of view and for a more complete bibliography.
[3] H. Fröhlich, Proc. Roy. Soc. (London) **A160**, 230 (1937); **A172**, 94 (1939); **A178**, 493 (1941); **A188**, 521 (1947); R. J. Seeger and E. Teller, Phys. Rev. **54**, 515 (1939); **56**, 352 (1939); H. Callen, Phys. Rev. **76**, 1394 (1940). F. Seitz, Phys. Rev. **76**, 1376 (1949).
[4] F. Seitz, Phys. Rev. **76**, 1376 (1949).
[5] H. Fröhlich, Proc. Roy. Soc. (London) **A188**, 532 (1947).
[6] W. R. Heller, Phys. Rev. **84**, 1130 (1951).
[7] W. Franz, Z. Physik **132**, 285 (1952).
[8] J. Yamashita and M. Watanabe, Repts. Inst. Sci. Tech. Univ. Tokyo **6**, III (1952).
[9] H. Kawamura, J. Phys. Soc. (Japan) **8**, 424 (1953).

of energy $\epsilon \approx 0$ are created, whereas one fast electron of energy $\epsilon \approx E$ disappears. The slow electron may either become a fast electron and create another pair of slow electrons by impact, or it may disappear from the conduction band by recombination. Breakdown occurs if the rate at which slow electrons become fast electrons is larger than the rate of recombination. Therefore, the breakdown theory is concerned with (a) the rate at which slow electrons become fast electrons, and (b) the rate of recombination.

An electron exchanges energy with the crystal lattice at an average rate $\langle d\epsilon/dt\rangle_L$, which is a function of its energy ϵ.[4] A typical curve of $\langle d\epsilon/dt\rangle_L$ *versus* ϵ is shown schematically in Fig. 2, case $F=0$. Except for electrons of very low energy $\epsilon \lesssim kT$, the electrons lose energy on the average by interaction with lattice vibrations. On the other hand, the electron gains (kinetic) energy from an applied field F at an average rate of $\langle d\epsilon/dt\rangle_F > 0$. The average rate of energy change from both field and lattice,

$$d\epsilon/dt = \langle d\epsilon/dt\rangle_L + \langle d\epsilon/dt\rangle_F, \qquad (1)$$

is shown schematically in Fig. 2 for three different fields. For sufficiently strong fields, $F \geq F^*$, we have $d\epsilon/dt \geq 0$ for all electron energies. For fields $F < F^*$, there exists an energy region where $d\epsilon/dt < 0$; this region will be referred to as "energy barrier" in what follows.

If the rate of energy change of each electron of energy ϵ would equal the average rate $d\epsilon/dt$, a slow electron could never become a fast electron, capable of forming a secondary electron, at a field $F < F^*$, and no breakdown could occur. Von Hippel identifies F^* with the breakdown field. Calculations of the breakdown field on this basis give a reasonable order of magnitude and predict an increase of the breakdown field with increasing temperature.

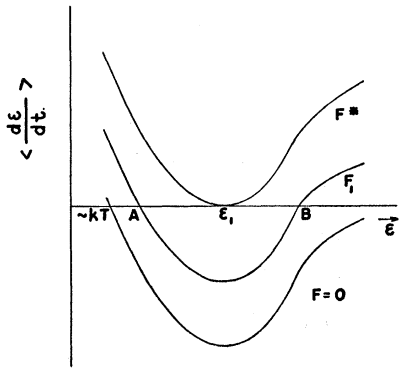

FIG. 2. Average rate of energy change $\langle d\epsilon/dt\rangle$ of an electron in the conduction band as function of its energy (schematically). $F=0$: without applied field; $\langle d\epsilon/dt\rangle = \langle d\epsilon/dt\rangle_L$; $F=F^*$: at von Hippel's breakdown field; the electrons of any energy gain energy on the average; $F_1 < F^*$: at a field less than von Hippel's breakdown field, there exists a range of energy $A \leq \epsilon \leq B$, where electrons lose energy on the average. This range will be referred to as an "energy barrier."

Experiments show that the breakdown field passes through a maximum with increasing temperature.[10] A possible explanation[11] of the decrease of the breakdown field with increasing temperature in the high-temperature region may be sought along the following lines. The energy change of an electron by interaction with the lattice is a statistical process, the elements of which are the absorption and emission, respectively, of a vibrational energy quantum $h\nu$ (phonon). At a sufficiently high temperature ($kT \gg h\nu$) and for electrons of energy $\epsilon \gg kT$, the rate of emission is larger than the rate of absorption; however, this difference is small as compared to either rate and arises from the "spontaneous emission." Hence, the change of energy of an electron by interaction with the lattice vibrations consists of large random fluctuations with a resultant small average energy. This is similar to the motion of an electron in a crystal in an electric field where there is a large random movement with a small superimposed directed movement. It is well known that the random movement leads

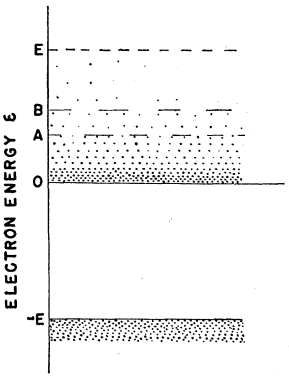

FIG. 3. Distribution of electrons (black dots) over the energy states of the crystal. Zero level of energy is the bottom of the electron conduction band. The top of the filled band lies at $-E$. The dotted line at $+E$ marks the electron energy necessary for production of secondary electrons by impact. In the range between A and B, the electrons lose energy on the average. This range is referred to as "energy barrier." A and B depend on the applied field.

to a diffusion current in case of a variation of concentration of electrons with position. Accordingly, in the case of a variation of electron concentration with energy (e.g., in the case of a Boltzmann distribution), the random fluctuation of energy leads to a diffusion current in energy space. Since there are more electrons of low energy than of high energy (see Fig. 3), the diffusion current is directed from low-energy values to high-energy values. We conclude that even in the case of $F \leq F^*$, a certain number of electrons may cross the "energy barrier" by diffusion processes.

It is obvious that the theory of electric breakdown which is based on "average electrons" does not take into account the crossing of the energy barrier by diffusion. The diffusion process enhances the number of fast electrons capable of producing secondaries and thus enhances electric breakdown. Hence one may suspect

[10] A. von Hippel and G. M. Lee, Phys. Rev. **59**, 824 (1941); A. M. Thomas and M. V. Griffith, J. Inst. Elect. Engrs. (London) **89**, 487 (1942).

[11] Another explanation has been proposed by H. Fröhlich, Proc. Roy. Soc. (London) **A188**, 521 (1947).

that the diffusion process, which increases with temperature, may be responsible for the observed decrease of the breakdown field at high temperatures.

By comparing the rates of crossings of the energy barrier due to the field and due to the diffusion, it is possible to modify the old breakdown theory[2] (which does not consider diffusion). This procedure, though only approximate, is attractive because it does not necessitate any assumptions about recombination processes. A detailed theory of breakdown must be based on knowledge of the distribution function of electrons over the energy states of the empty band and necessitates detailed assumption regarding their recombination with holes in the filled band.[6,7]

RATE OF FLOW OF ELECTRONS IN WAVE VECTOR SPACE

Consider the number of electrons per unit time which change from energy less than ϵ to larger than ϵ. This number represents a flow rate $S(\epsilon)$ in wave vector space through the surface $\epsilon=$const. Let $0\leq f(\epsilon)\leq 1$ be the probability that an energy state ϵ is occupied by an electron. In what follows we shall express $S(\epsilon)$ in terms of the distribution function $f(\epsilon)$ of the electrons. The calculation will be based on the following assumptions:

(1) Spherically symmetrical energy surfaces in wave vector space

$$\epsilon=h^2k^2/2m \tag{2}$$

(h Planck's constant, k wave vector of the electron, m effective electron mass). For such surfaces the density of energy states is[12]

$$D(\epsilon)=2(2\pi/h^3)(2m)^{\frac{3}{2}}\epsilon^{\frac{1}{2}}. \tag{3}$$

(2) The distribution function $f(\epsilon)$ can be expanded in terms of Legendre polynomials with the field direction as axis. All terms except the first will be neglected. Thus we set

$$f(k,\vartheta)=f_0(\epsilon)+f_1(\epsilon)k_x, \tag{4}$$

where $k_x=|k|\cos\vartheta$ is the wave vector component in field direction.

(3) A detailed analysis of the scattering process and of the influence of the field on the electron distribution shows that at sufficiently weak fields

$$f_1\sim-\tau(eF/m)h(df_0/d\epsilon), \tag{5}$$

where

$$1/\tau=-(1/k_x)\langle dk_x/dt\rangle_L \tag{6}$$

is the "transport collision frequency." τ is a function of energy in a manner which depends on details of the scattering mechanism of electrons and lattice vibrations. In what follows, we shall assume that Eq. (5) is valid for field strengths up to the breakdown field.

(4) The electrons interact with lattice vibrations of a single frequency ν (Einstein distribution), which are assumed to be thermally fully excited; i.e., $n=kT/h\nu\gg 1$ (k Boltzmann constant).

(5) Collisions between electrons are neglected (see Heller[6]).

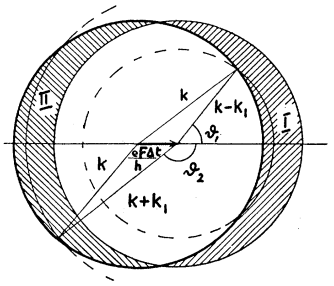

FIG. 4. A shift of the electron distribution by $eF\Delta t/h$ in the k_x direction causes the electrons of wave vectors in area I to flow out of the energy surface $\epsilon=(h^2k^2/2m)$ (thickly drawn circle) and the electrons of wave vectors in area II to flow into the energy surface.

There are three contributions to the flow of electrons in k-space $S(\epsilon,F)$ through the energy surface $\epsilon=$constant:

$$S=S_F+S_L+S_D. \tag{7}$$

S_F arises from the change of wave vector with time due to the electric field

$$(dk_x/dt)_F=eF/h. \tag{8}$$

We have (see Fig. 4)[13]

$$S_F dt=4\pi\int_0^{eFdt/h}\left\{\int_0^{\vartheta_1(k_1)} f(k-k_1,\vartheta)(k-k_1)^2\sin\vartheta d\vartheta - \int_{\vartheta_2(k_2)}^{\pi} f(k+k_1,\vartheta)(k+k_1)^2\sin\vartheta d\vartheta\right\}dk_1. \tag{9}$$

Considering a sufficiently short time interval dt so that $(eF/h)dt\ll k$, we obtain from Fig. 4: $\cos\vartheta_1=-\cos\vartheta_2\cdot k_1/(eF/hdt)$. If we introduce Eq. (4) into Eq. (9), the terms containing f_0 cancel and we obtain

$$S_F=(8\pi/3)(eF/h^4)(2m)^{\frac{3}{2}}f_1(\epsilon)\epsilon^{\frac{1}{2}}. \tag{10}$$

Combining Eqs. (10), (3), and (5), we have

$$S_F=-\tfrac{2}{3}D(\epsilon)(eF)^2\epsilon(df_0/d\epsilon)/m. \tag{11}$$

S_L arises from the average change of energy due to collisions with lattice vibrations,

$$S_L=Df_0\langle d\epsilon/dt\rangle_L. \tag{12}$$

In the temperature region where $n=kT/h\nu\gg 1$, and for energies $\epsilon\gg h\nu$, the occurence of the absorption and emission of vibrational quanta $h\nu$ is nearly equally frequent. The energy change $\langle d\epsilon/dt\rangle_L$ results then only from the spontaneous emission and is on the average $-h\nu/(2n+1)$ per collision, whereas the random energy change per collision is $\pm h\nu$. Thus,

$$\langle d\epsilon/dt\rangle_L=-\nu_c h\nu/(2n+1)\approx-\nu_c(h\nu)^2/2kT, \tag{13}$$

where ν_c is the collision frequency. S_D results from the random changes $\pm h\nu$ of electron energy per collision; if

[12] $D(\epsilon)$ includes a factor 2 from the two spin possibilities.

[13] A factor 2 accounting for the two spin possibilities is included.

we neglect the changes of $D(\epsilon)$ with ϵ in the range $\epsilon-h\nu$, $\epsilon+h\nu$, we obtain:

$$S_D=-(\nu_c/2)\int_0^{h\nu} D[f_0(\epsilon+\epsilon_1)-f_0(\epsilon-\epsilon_1)]d\epsilon_1$$
$$\approx-(\nu_c/2)D(df_0/d\epsilon)(h\nu)^2$$
$$=D(df_0/d\epsilon)\mathrm{k}T\langle d\epsilon/dt\rangle_L. \quad (14)$$

It should be noted that Eqs. (13) and (14) are valid only if $h\nu\ll\epsilon$.

By substituting Eqs. (11), (12), and (14) into Eq. (8), we obtain:[5]

$$S(\epsilon,F)=Df_0\langle d\epsilon/dt\rangle_L$$
$$+Ddf_0/d\epsilon[\mathrm{k}T\langle d\epsilon/dt\rangle_L-(2/3)\tau\epsilon(eF)^2/m]. \quad (15)$$

$\langle d\epsilon/dt\rangle_L$ and τ are functions of ϵ, which may be derived from considerations of the type of interaction of electrons with lattice vibrations. For the purposes of this paper, we need not specify these functions in detail, except that $\langle d\epsilon/dt\rangle_L$ is negative for $\epsilon\geq kT$, and that there is a maximum of $-(\langle d\epsilon/dt\rangle_L/\tau)$ at an energy ϵ_1, (see Fig. 2).[14]

COMPARISON WITH THE OLD BREAKDOWN THEORY

According to Eqs. (16) and (8), the average change of wave vector component in the x direction is

$$dk_x/dt=\langle dk_x/dt\rangle_F+\langle dk_x/dt\rangle_L=(eF/h)-(k_x/\tau). \quad (16)$$

The ordinary breakdown theory[4,15] considers an "average electron," defined by having a wave vector component in the x-direction which does not change, on the average, because of the field or lattice vibrations, i.e., $dk_x/dt=0$. Hence the x-component of wave vector of the average electron is

$$k_x=(eF/h)\tau. \quad (17)$$

The average rate of change of energy due to the field is obtained from Eqs. (2) and (8):

$$\langle d\epsilon/dt\rangle_F=(d\epsilon/dk_x)\langle dk_x/dt\rangle_F=hk_x(eF/m). \quad (18)$$

From Eqs. (18) and (17), the average rate of change of energy of the "average electron" is obtained:

$$d\epsilon/dt=\langle d\epsilon/dt\rangle_L+[(eF)^2/m]\tau. \quad (19)$$

According to von Hippel,[2] the breakdown field F^* is determined by

$$d\epsilon/dt=\langle d\epsilon/dt\rangle_L+[(eF^*)^2/m]\tau=0, \quad (20)$$

where the right side is to be taken at its maximum value (the corresponding energy will be designated by ϵ_1; see Fig. 2).

If we take the flow rate $S_{F=F^*}$ at $\epsilon=\epsilon_1$ as determining the breakdown, then the inclusion of the diffusion flow rate S_D would lead to a breakdown field G^*, defined by

$$S_{F=F^*}=S_{F=G^*}+S_D, \quad (21)$$

with all values taken at $\epsilon=\epsilon_1$. Equations (11) and (14) indicate that $S_F\rightarrow S_F+S_D$ by the substitution $(eF)^2/m=(eF)^2/m-\frac{3}{2}(\mathrm{k}T/\epsilon)\langle d\epsilon/dt\rangle_L/\tau$. Hence the breakdown field F^* as given by the ordinary breakdown theory [Eq. (20)] should be corrected to give the breakdown field G^* defined by Eq. (21) by setting:

$$G^{*2}=F^{*2}+\tfrac{3}{2}(m/e^2)\langle d\epsilon/dt\rangle_L/(\mathrm{k}T/\epsilon_1)$$
$$=F^{*2}[1-\tfrac{3}{2}(\mathrm{k}T/\epsilon_1)]. \quad (22)$$

If F^{*2} increases proportionally to T, G^* has a maximum $F^*/2$ at the temperature $T_m=\epsilon_1/(3\mathrm{k})$.

In the case of interaction of electrons with the optical modes of lattice vibrations in ionic crystals, we have $\epsilon_1\sim0.1-0.2$ ev (e.g., most of the alkali halides) corresponding to $T_m\sim400-800°\mathrm{K}$. Therefore, at ordinary temperatures, the term $\frac{3}{2}\mathrm{k}T/\epsilon_1$ is not negligible compared to unity in the case of interaction with the optical modes of lattice vibrations. In the case of interaction with the acoustical modes of lattice vibrations, ϵ_1 is of the order of an electronvolt. Hence the term $\frac{3}{2}\mathrm{k}T/\epsilon_1$ should be negligible compared to unity in the case of interaction with the acoustical modes, in agreement with the estimate by Seitz.[3]

In ionic crystals either the interaction with the optical modes of lattice vibrations or that with the acoustical modes of lattice vibrations may be the cause of the "energy barrier," determining the breakdown field.[4] It is usually assumed[15] that the interaction with the optical modes determines the breakdown field, though this is not necessarily the case.

GENERAL TREATMENT OF BREAKDOWN BASED ON THE DISTRIBUTION FUNCTION

The breakdown criterion of the simple theory is unsatisfactory because it does not compare the rate of generation of secondary electrons with the rate of their recombination. The breakdown criterion should be based on the condition that no stationary electron distribution can be maintained for fields $F\geq H$ (where H is the breakdown field). In order to obtain a quantitative formulation, assumptions about (a) the ionization process, and (b) the recombination process, have to be made.

With regard to the ionization process, we shall assume that each electron creates a secondary electron and loses its own energy as soon as it acquires the ionization energy E.[16] This means that (1) fast electrons disappear at a rate $S(E)$, (2) slow electrons are created at a rate $2S(E)$, and (3) that $f(\epsilon)=0$ for $\epsilon>E$.

With regard to the recombination process, we shall assume:

[14] Since $\tau\approx1/\nu_c$ and since ν_c is proportional to $\langle d\epsilon/dt\rangle_L$, $\langle d\epsilon/dt\rangle_L/\tau$ is essentially the function $\langle d\epsilon/dt\rangle_L^2$.

[15] H. Callen, Phys. Rev. **76**, 1394 (1949).

[16] As is shown by Heller (see reference 6) the delay of the distribution function for energies $\epsilon>E$ is extremely rapid.

(1) Only "slow electrons" of energy $\epsilon<\epsilon_0$ recombine. This assumption seems justified since only slow electrons remain sufficiently long in the neighborhood of recombination centers to recombine. We shall choose $\epsilon_0 \gtrsim kT \gg h\nu$, so that $\langle d\epsilon/dt\rangle_L<0$ for $\epsilon \geq \epsilon_0$ (see Fig. 2, case $F=0$).

(2) The rate of recombination is a given function R of $f(\epsilon_0)$, which we shall not specify closer. At first sight it appears preferable to consider the recombination as a function of the number of "slow electrons"

$$N=\int_0^{\epsilon_0} Df_0(\epsilon)d\epsilon.$$

However, if ϵ_0 is of the order kT, $f_0(\epsilon)$ will be nearly a constant in the range $\epsilon_0 \leq \epsilon \leq E_0$ and N will be proportional to $f_0(\epsilon_0)$.

Since there is no recombination in the range $\epsilon_0 \leq \epsilon \leq E$, we have $f_0(\epsilon)=$constant in this range in a stationary state, and we obtain from Eq. (15):

$$f_0(\epsilon)=f_0(E)\exp\left(+\int_\epsilon^E p d\epsilon\right) - S\left[\int_\epsilon^E p\left\{\frac{\exp\left(-\int_\epsilon^E p du\right)}{D\langle d\epsilon/dt\rangle_L}\right\}d\epsilon\right] \times\left[\exp\left(+\int_\epsilon^E p d\epsilon\right)\right], \quad (23)$$

where

$$p=[kT-(2/3)(\tau\epsilon/\langle d\epsilon/dt\rangle_L)(eF)^2/m]^{-1}. \quad (24)$$

If one recalls that $\langle d\epsilon/dt\rangle_L<0$ for the energy range considered, and that $0<p\leq(kT)^{-1}$, Eq. (23) may be interpreted as follows: The first term on the right hand side is a generalized Boltzmann distribution with a "temperature" p/k depending on energy and field as expressed by Eq. (24). The second term is the modification of the Boltzmann distribution by the flow S. It will be shown in the Appendix, that $f_0(E)$ is proportional to S, the proportionality constant being a function of the applied field:

$$f_0(E)=a(F)S. \quad (25)$$

If we introduce the Eq. (25) into the Eq. (23), written for $\epsilon=\epsilon_0$, we obtain

$$b(F)=S/f_0(\epsilon_0) = \left\{a(F)-\int_{\epsilon_0}^E p\left[\frac{\exp\left(-\int_\epsilon^E p du\right)}{D\langle d\epsilon/dt\rangle_L}\right]d\epsilon\right\}^{-1} \times\exp\left(-\int_{\epsilon_0}^E p d\epsilon\right). \quad (26)$$

Since $a(F)>0$ (see Appendix) and $\langle d\epsilon/dt\rangle<0$ (for $\epsilon\gtrsim\epsilon_0$), we have $b(F)>0$. For $F\to\infty$, we have $p\to 0$ [Eq. (24)] and $a\to 0$ (Appendix), and therefore $b(F)\to\infty$.

In order to obtain the breakdown field, we have to compare the flow $S=bf(\epsilon_0)$ with the recombination $Rf_0(\epsilon_0)$. At fields below breakdown, the value f_0 is determined by the balance between production of secondaries and recombination:

$$b(F)f_0(\epsilon_0)=R[f_0(\epsilon_0)]. \quad (27)$$

The breakdown field H is determined by the condition that no solution $f_0(\epsilon_0)\leq 1$ of the Eq. (27) exists for $F\geq H$, whereas such a solution exists for $F<H$. Any extension of the above theory requires more detailed knowledge of the recombination process.

RELATION TO THE THEORY OF THERMIONIC EMISSION

It will be noticed that for $F=0$, Eq. (26) leads to a finite rate of ionization S_0 and Eq. (23) gives then an electron distribution which differs from that of thermal equilibrium. The reason for the failure of Eq. (23) in the case of thermal equilibrium ($F=0$) is the following: Our treatment was based on the creation of two slow electrons and the disappearance of one fast electron in the process of the ionization by impact (see Fig. 1). The reverse process, disappearance of two electrons and simultaneous creation of one fast electron, was neglected. However, in thermal equilibrium the rate of each process equals that of its reversed process. Hence we cannot keep one process and neglect its reverse process in thermal equilibrium.

A finite flow rate S in the case of $F=0$ and the corresponding deviation from thermal equilibrium distribution of electrons are of interest for the theory of thermionic emission. The thermionic emission current results from the few electrons which have by thermal agitation acquired sufficient energy to overcome the potential wall at the surface of the solid. In the theory of breakdown fast electrons are removed by ionization processes. In the theory of thermionic emission fast electrons are removed irreversibly over the surface barrier of the solid (Fig. 5).

The calculations of thermionic emission are usually based on the assumption of a thermal equilibrium distribution of electrons within the solid. Clearly this assumption is justified only if the rate at which fast electrons leave the solid is small compared to the rate at which fast electrons can be replenished in the boundary of the solid.

Fast electrons can be replenished in the boundary by slow electrons gaining energy near the surface of the cathode and by fast electrons from the bulk of the solid entering the boundary zone (Fig. 5). A detailed discussion of these electron currents would require the consideration of a distribution function $f_0(\epsilon,x)$, which

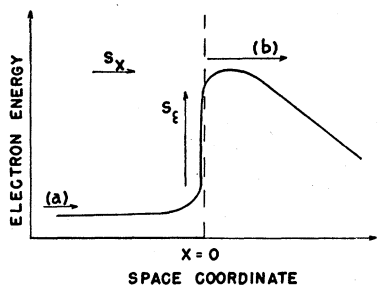

FIG. 5. Energy profile along a line perpendicular to the surface ($x=0$) of a cathode with a field applied. Electrons of low energy are supplied from the bulk of the cathode (arrow *a*); the thermionic emission removes electrons of high energy from the surface of the cathode (arrow *b*). These electrons are replenished by slow electrons gaining energy (vertical arrow S_ϵ) or fast electrons moving toward the boundary (horizontal arrow S_x).

depends on both, energy ϵ and position.[17] From the distribution function a flow vector $\mathbf{S}(\epsilon,x)$ can be derived. In a stationary state one has

$$(\partial S_\epsilon/\partial\epsilon)+(\partial S_\psi/\partial x)=0, \tag{28}$$

where S_ϵ and S_x are, respectively, vertical and horizontal components of the flow vector in a diagram such as in Fig. 5. The component S_x can be represented as the sum of a spacial field current and of a spacial diffusion current. If the expressions for S_ϵ [Eq. (15)] and S_x are introduced into Eq. (28), a partial differential equation for the function $f_0(\epsilon,x)$ is obtained. Boundary conditions are: The quasithermal equilibrium distribution, modified by a field, in the bulk of the cathode ($x=-\infty$) and the thermionic emission current resulting from $f_0(\epsilon,0)$ at the surface of the cathode.[18]

In absence of a solution of the two-dimensional problem mentioned above, we shall attempt to estimate an upper limit of thermionic emission, imposed by the rate at which slow electrons can gain energy near the surface of a cathode. The emission current is

$$I=e\int_{-\infty}^{0} S_\epsilon(E,x)dx. \tag{29}$$

In order to estimate this integral we shall replace the gradual decrease of $S_\epsilon(E,x)$ toward the inside of the cathode by a step function (see Fig. 6); i.e.,

$$S_\epsilon(E,x)=S, \quad \text{for} \quad -L\leq x\leq 0;$$
$$S_\epsilon(E,x)=0, \quad \text{for} \quad x<-L.$$

We have then

$$I\sim eLS. \tag{30}$$

[17] The component S_ϵ cannot be considered as being independent of x any more.

[18] In studying emission by the tunnel effect, the boundary condition at $x=0$ has to take into account the emission of electrons through the surface barrier by tunneling.

In the case of a negligible electric field,[19] $F=0$ in the cathode, the value of L can be obtained by considering the diffusion of electrons of energy $\epsilon>E$ toward the surface in a layer of thickness L with a source density S of electrons of energy $\epsilon\geq E$. The continuity equation for the concentration,

$$n=\int_\epsilon^\infty f_0(\epsilon,x)D(\epsilon)d\epsilon, \tag{31}$$

of electrons of energy $\epsilon\geq E$ reads

$$-\mathfrak{D}(d^2n/dx^2)e=Se, \tag{32}$$

where $\mathfrak{D}$ is the (average) diffusion constant of electrons of energy $\epsilon\geq E$. The solution of Eq. (32) with the boundary conditions $n=n^*$ (n^* is the thermal equilibrium value) at $x=-L$ and $n=n_0$ at $x=0$ reads

$$n=n_0-(n^*-n_0)x/L-(S/2\mathfrak{D})x(x+L). \tag{33}$$

The condition that

$$dn/dx=0 \quad \text{at} \quad x=-L \tag{34}$$

yields

$$L=[(n^*-n_0)2\mathfrak{D}/S]^{\frac{1}{2}}. \tag{35}$$

Therefore, from the Eqs. (30) and (35), we obtain

$$I<e(2n^*\mathfrak{D}S)^{\frac{1}{2}}. \tag{36}$$

By writing Eq. (28) for $\epsilon=E$ and $F=0$,[20] we obtain

$$f_0(E)\approx f_0(\epsilon_0)\exp[-(E-\epsilon_0)/kT]+S/\langle d\epsilon/dt\rangle_L. \tag{37}$$

It will be noticed that $S(>0)$ is less than the value,

$$S_M=-f_0(\epsilon_0)\exp[-(E-\epsilon_0)/kT]\langle d\epsilon/dt\rangle_L D, \tag{38}$$

obtained from Eq. (37) by setting $f_0(E)=0$. Therefore, from Eq. (36) with (38), $f_0(\epsilon_0)\approx\exp-[(\epsilon_0-\zeta)/kT]$ and $n^*=D(E)kT\exp[-(E-\zeta)/kT]$ [where $D(E)$ is

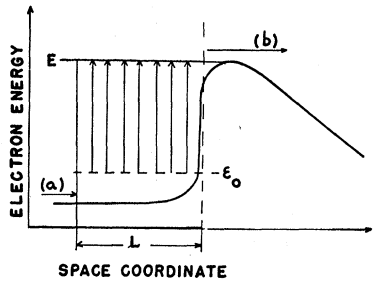

FIG. 6. Approximation of the electron flow for sake of obtaining an order of magnitude estimate of its effect on thermionic emission.

[19] The charges induced in semiconducting cathodes by the applied voltage may extend over many interatomic distances into the cathode [N. D. Morgulis, J. Phys. (U.S.S.R.) **11**, 67 (1947)], and the neglection of the electric field in the cathode may not be justified then; the considerations of the text apply to a metal cathode.

[20] Use of Eq. (28) implies that $\partial S_x/\partial x$ for $\epsilon<E$ is negligible *versus* $\partial S_\epsilon/\partial\epsilon$; it follows then that $S_\epsilon(\epsilon,x)$ is independent of ϵ, and that Eq. (28) is valid.

defined by Eq. (3)]; we obtain

$$I<AT^2\exp[-(E-\zeta)/kT]\times[(16\mathfrak{D}mE\langle d\epsilon/dt\rangle_L)/(kT)^3]^{\frac{1}{2}}, \quad (39)$$

where

$$A=4\pi mek^2/h^3 \quad (40)$$

is the Richardson constant, and $E-\zeta$ is the "work function." Therefore, deviations from the Richardson equation should be expected, if

$$(16\mathfrak{D}mE\langle d\epsilon/dt\rangle_L)/(kT)^3\lesssim 1. \quad (41)$$

Condition (41) can be transformed into

$$Ehv/(kT)^2\lesssim(3/16)^{\frac{1}{2}} \quad (42)$$

by using the values of Eq. (13) for $\langle d\epsilon/dt\rangle_L$ and[21]

$$\mathfrak{D}\approx 2E/(3mv_c). \quad (43)$$

Since our considerations were based on the assumption $E\gg kT\gg h\nu$, it is not obvious whether the condition (42) is fulfilled or not.

The simple estimate presented above was based on the assumption that all electrons of energy $\epsilon>E$ have the same ("average") velocity. No conclusion can be drawn therefore from these considerations about the energy distribution of the emitted electrons. Deviations from a Boltzmann distribution of emitted electrons have been observed.[22]

In the above discussion we have neglected the influence of the electric field in the cathode on the emission. At high fields, the effect of the electric field on the factor S of Eq. (30) is stronger than the effect of the electric field on the factor L. Therefore, it may be permissible for an order of magnitude estimate of the field dependence to consider L in Eq. (30) as a constant.[23] Using the value of S from Eq. (23) (written for $\epsilon=E$) and expressing $f_0(E)$ by $I\approx AT^2f_0(E)$,[24] we obtain from Eq. (30):

$$I=\frac{AT^{*2}f_0(\epsilon_0)\exp-(E-\epsilon_0)/kT}{1+d}c, \quad (44)$$

where c and d are numerical values, given by

$$c=\exp\left[-\int_{\epsilon_0}^{E}\langle p-(kT)^{-1}\rangle d\epsilon\right]\geq 1, \quad (45)$$

$$d=-AT^2\int_{\epsilon_0}^{E}\left[\frac{p\exp\left(-\int_{\epsilon_0}^{E}pdu\right)}{LeD\langle d\epsilon/dt\rangle_L}\right]d\epsilon>0. \quad (46)$$

[21] According to E. H. Kennard [*Kinetic Theory of Gases* (McGraw-Hill Book Company, Inc., New York, 1938), p. 188] the diffusion current of gas particles of concentration gradient dn/dx and of average velocity $\bar{v}$ is $\frac{1}{3}(\bar{v})^2\tau(dn/dx)$. Accordingly, the diffusion coefficient is $\mathfrak{D}=\frac{1}{3}(\bar{v})^2\tau$. We are dealing with electrons of energy $\geq E$, velocity $\bar{v}\approx(2E/m)^{\frac{1}{2}}$, and collision time $1/\nu_c$. Insertion of these values into Kennard's formula gives the diffusion coefficient of Eq. (43).

[22] W. B. Nottingham, Phys. Rev. **49**, 78 (1936).

[23] In case of strong fields the divergence of the field current has to be added to Eq. (32), and Eq. (33) is then no longer valid.

[24] Strictly, the temperature T of the factor T^2 should be replaced by $T^*=(L/k)f_0(\epsilon)/[df_0(\epsilon)/d\epsilon]$ at $\epsilon=E$.

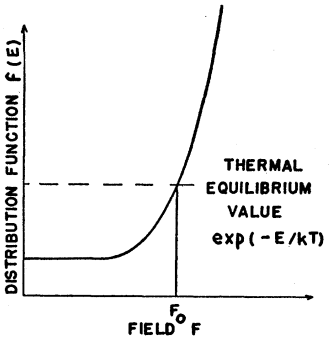

FIG. 7. Probability of occupation by an electron for an energy state near the top of the surface barrier of a cathode from which a saturation current is drawn as function of the field in the cathode. The thermionic emission current is proportional to $f(E)$.

In the case of $F=0$, we have

$$c_0=1, \quad (47)$$

and

$$d_0=-AT^{*2}\int_0^E\frac{\exp[-(E-\epsilon)/kT]}{kTDLe\langle d\epsilon/dt\rangle_L}d\epsilon \approx-(AT^{*2}/Le)[1/(D\langle d\epsilon/dt\rangle_L)], \quad (48)$$

where the values of L, D and $d\epsilon/dt$ are to be taken at $\epsilon=E$. With increasing F, d increases and c decreases. However, the change of d and c with F becomes noticeable only when $p(\epsilon)$ differs appreciably from $1/kT$. If E is larger than the peak ϵ_1 of the friction barrier (see Fig. 2), the difference between $1/kT$ and $p(\epsilon,F)$ will be largest at $\epsilon=\epsilon_1$. Thus the appreciable change of c and d commences at a field F_0 which is determined by the value of $(1/kT)-p_{\epsilon=\epsilon_1}$, but does not depend critically on the height of the surface barrier E (if $E\gg\epsilon_1$). For F larger than F_0, the thermionic emission rises rapidly, because the concentration of electrons of energy E rises rapidly, as shown schematically in Fig. 7.

In the literature, the increase of the thermionic emission current at strong fields is usually attributed to a decrease of the surface barrier E with the field[25] or to tunneling of electrons through the surface barrier. The above considerations indicate an additional effect, caused by deviations of the electron distribution from thermal equilibrium at fields of the order of F_0.

CONCLUSIONS

We have defined at various points of this paper characteristic fields which are of importance for breakdown (F^*,H) and for a rapid increase in thermionic emission with the field (F_0). The question arises whether there exists a relation between these fields.

F^* is the field at which the electron of any energy gains energy on the average as required by von Hippel's breakdown condition, Eq. (20). If we introduce F^* into the Eq. (26), a flow rate $S(F^*)$ would be obtained. Depending on the assumption made for the recombination process, the flow rate $S(H)$, corresponding to the

[25] W. Schottky, Z. Physik **15**, 872 (1914); H. Morgulis, J. Phys. (U.S.S.R.) **11**, 67 (1947); D. A. Wright, Proc. Phys. Soc. (London) **B65**, 134 (1952); K. Lehovec, Phys. Rev. **85**, 705 (1952).

breakdown field H, obtained from Eq. (27), may be smaller or larger than the flow rate $S(F^*)$. Thus, there is no direct comparison possible between the breakdown field as obtained by the simple theory and by the refined theory. In fact, there is little quantitative justification for the breakdown criterion as used in the simple theory.

However, a certain relation between H and F^* is indicated from the following more qualitative considerations. A closer analysis of Eq. (26) shows that S depends on F mainly through the parameter p. At small F, $p\approx 1/kT$. At a given F, p deviates most strongly from $1/kT$ at the position $\epsilon=\epsilon_1$; i.e., the top of the friction barrier. Thus, $S(F)$ becomes strongly dependent on F, if $(\lambda/kT)-p(F,\epsilon_1)=\Delta$ differs appreciably from $1/kT$. For von Hippel's breakdown field F^* [Eq. (20)]: $\Delta=(kT)^{-1}[1+3kT/(2\epsilon_1)]^{-1}$.

Similar considerations apply to a discussion of the field F_0, at which thermionic emission rises rapidly, since this rise depends on F again through the parameter p. Since the strongest deviation of p from $1/kT$ occurs at $\epsilon=\epsilon_1$ (if $E\gg\epsilon_1$) and not at E, the field F_0 characteristic for a rapid rise of thermionic emission is rather insensitive to the height of the surface barrier. Similarly, the field H at which breakdown occurs should be rather insensitive to the ionization energy in accordance with the wellknown "low-energy criterion." In conclusion, thermionic emission should rise rapidly with the field, at fields of the same order of magnitude as the breakdown field, if both the ionization energy and the height of the surface barrier are larger than the energy corresponding to the top of the friction barrier.

In this paper, only a stationary state has been considered. Since most of the breakdown experiments are carried out under pulse conditions in order to avoid heating as much as possible, a stationary state may not be reached in these experiments and breakdown field may then depend on pulse duration as discussed by Franz.[7]

APPENDIX

Relation between $f_0(E)$ and S

Consider the flow of electrons through the energy surface $\epsilon=E$. Since there are no electrons of energy $\epsilon>E$, we have no flow from the outside through the energy surface $\epsilon=E$. Thus the flow through the energy surface $\epsilon=E$ is only due to (a) energy gain of electrons in the field, (b) absorption of phonons by electrons (see Fig. 4):

$$S_F dt=4\pi\int_0^{eFdt/h}\int_0^{\vartheta_1(k)} f(k-k_1,\vartheta)(k-k_1)^2\sin\vartheta d\vartheta dk_1$$
$$=2\pi[f_0(E)(2mE/h^2)+\tfrac{2}{3}f_1(E)(2mE/h^2)^{\frac{1}{2}}]eFdt/h. \quad \text{(A-1)}$$

Expressing $f_1(E)$ by $df_0/d\epsilon/_{\epsilon=E}$, and using Eq. (5), we obtain from Eq. (23),

$$(df_0/d\epsilon)_{\epsilon=E}=p[(S/D)\langle d\epsilon/dt\rangle_L-f_0(E)]. \quad \text{(A-2)}$$

Combining the Eqs. (A-1), (5), and (A-2) and replacing D by the value given in Eq. (3), we have

$$S_F=f_0(E)[4\pi mEeF/h^3+(4/3)(p\tau e^2F^2/m)(2mE/h^2)^{\frac{1}{2}}]-(SpE\tau e^2F^2)/(3m\langle d\epsilon/dt\rangle_L). \quad \text{(A-3)}$$

The right hand side is to be taken at $\epsilon=E$. The flow of electrons through the energy surface $\epsilon=E$, caused by the absorption of phonons, is

$$S_A=(\nu_c/2)[h\nu Df_0(E)]. \quad \text{(A-4)}$$

We may express ν_c by $\langle d\epsilon/dt\rangle_L$, using Eq. (13), and obtain from the Eqs. (A-3) and (A-4):

$$S=f_0(E)\{4\pi mEeF/h^3+(4/3)(p\tau e^2F^2/m)(2mE/h^2)^{\frac{1}{2}}-(\mathrm{k}T\langle d\epsilon/dt\rangle_L D)/(h\nu)\}-(SpE\tau e^2F^2)/(3m\langle d\epsilon/dt\rangle_L). \quad \text{(A-5)}$$

Therefore,

$$a(F)=\frac{f_0(E)}{S}=\frac{[1+(pE\tau e^2F^2)/(3m\langle d\epsilon/dt\rangle_L)]}{[4\pi mEeF/h^3+(4/3)(p\tau e^2F^2/m)(2mE/h^2)^{\frac{1}{2}}-(\mathrm{k}T\langle d\epsilon/dt\rangle_L D/(h\nu)]}. \quad \text{(A-6)}$$

The values of $a(F)$ for the two boundary cases $F=0$ and $F\to\infty$ are as follows: In the case $F=0$, one has $p=1/kT$ and

$$a(0)=-h\nu/(\mathrm{k}T\langle d\epsilon/dt\rangle_L D)_{\epsilon=E}, \quad \text{(A-7)}$$

and in the case $F\to\infty$, one has $p\approx-\frac{3}{2}m\langle d\epsilon/dt\rangle_L/\tau Ee^2F^2$, and Eq. (A-6) becomes

$$a(\alpha)\approx h^3/(8\pi mEeF). \quad \text{(A-8)}$$

ACKNOWLEDGMENT

The author is indebted to Miss Alma Marcus and Mr. Kurt Schoeni for helpful discussions.