SOVIET PHYSICS USPEKHI VOLUME 5, NUMBER 3 NOVEMBER-DECEMBER 1962

TUNNEL DIODES

V. I. FISTUL' and N. Z. SHVARTS

Usp. Fiz. Nauk **77**, 109-160 (May, 1962)

CONTENTS

INTRODUCTION

J. I. Frenkel and A. F. Ioffe, examining the nature of rectification on the boundary between a metal and semiconductor, arrived at a hypothesis that the mechanism of this phenomenon may be the electron tunnel effect. This idea was advanced by them in 1932 [1]. It could not be verified experimentally at that time, owing to the low level of semiconductor technology.

It took 25 years of vigorous development of physics and semiconductor technology for this idea of Frenkel and Ioffe to come to fruition in the tunnel diode, discovered by the Japanese scientist Esaki [5] in 1958.

The tunnel diode has attracted much attention from specialists in various fields. Its quantum-mechanical nature allows such diodes to be used for various very precise physical researches. The use of tunnel diodes affords new possibilities for the generation, amplification, mixing, and detection of electromagnetic oscillations.

The operating speed of the device makes it a very valuable computer element.

Many experimental and theoretical papers dealing with the physics and technology of tunnel diodes have been published by now in various countries.

The present article is a review of the papers on tunnel diodes published up to July 1961. Papers dealing with computer applications of tunnel diodes are not included; their number is large enough to warrant a separate review. A rather exhaustive bibliography of the papers dealing with this subject can be found in [129].

In preparing the present review, the authors paid principal attention to the physics of the tunnel diode and to its high-frequency applications in radio engineering.

1. OPERATING PRINCIPLE OF THE TUNNEL DIODE

Until recently the physics and technology of semiconductor diodes was limited to crystals in which the impurity concentration did not exceed 10^{16}—10^{17} cm^{-3}. Accordingly, the concentration of the free carriers in these crystals was of the same order of magnitude or even less. The corresponding location of the Fermi level and of the band boundaries in the semiconductors had the form shown in Fig. 1. The energy structure of the p-n junction (under thermodynamic equilibrium conditions) in such crystals is shown in Fig. 2.

A p-n junction which is homogeneous in structure and has a sharply delineated impurity distribution (alloyed junction) is given, as is well known [2], by the expression

$$L = 1{,}05 \cdot 10^6 \left[\frac{\varepsilon (U_k - U)}{2\pi q} \frac{n_d + n_a}{n_d n_a} \right]^{1/2} \text{[cm]}, \qquad (1.1)$$

where ϵ —dielectric constant, q —electron charge, n_d and n_a —densities of the ionized donors and acceptors on both sides of the p-n junction, U_k —contact potential difference on the junction, and U —external voltage applied to the junction.

If the impurity density is 10^{16}—10^{17} cm^{-3}, the thickness of the p-n junction in the germanium is approximately 10^{-4}—10^{-5} cm. The electric field $\mathscr{E}$ in such a junction is likewise small in the absence of bias. For germanium, for example, it amounts to

$$\mathscr{E} \cong \frac{U_k}{L} = \sim 10^3 - 10^4 \text{ [V/cm]}.$$

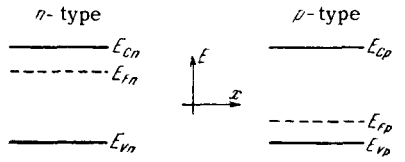

FIG. 1. Position of Fermi level in weakly-doped semiconductor.

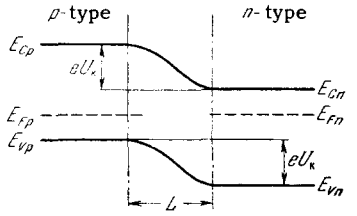

FIG. 2. Energy band scheme in p-n junction of ordinary semiconductor diode.

The p-n junction in an ordinary diode is thus characterized by rather large thickness and relatively low field intensity. In this case electrons can enter the free zone, i.e., overcome a certain energy barrier, only if additional energy is imparted to them. Large fields can be produced in an ordinary p-n junction only by large voltages. A strong field ($\mathcal{E} \sim 10^5$ V/cm) produces additional carriers, the number of which increases appreciably with increasing field. Many investigations, the principal among which are reported in [3,4], have disclosed that the most probable mechanism for this increase is impact ionization. Along with causing impact ionization, a strong field may directly extract the carriers from the filled band into the free one. Such an extraction is similar in nature to cold emission of electrons from a metal. This mechanism was most clearly manifest in the tunnel diodes, which were discovered in 1958 [5].

The distinguishing feature of the semiconductors used in tunnel diodes is an impurity content of 10^{18}—10^{20} cm^{-3}. The positions of the band boundaries and the Fermi level in such semiconductors are shown in Fig. 3, while the band structure under equilibrium conditions in a p-n junction produced in such a semiconductor is shown in Fig. 4. The thickness of the p-n junction decreases in this case, as follows from (1.1), to 100—200 Å. The field intensity in the junction increases sharply to 10^5—10^6 V/cm. Under these conditions the electron in the conduction band of the n-region of the crystal needs no additional energy to overcome the energy barrier. It can find its way to the other side of the p-n junction, in the valence band, by virtue of the so-called tunneling, due to the tunnel effect [6], whereby it is at the same energy level on both sides of the barrier. It can be assumed here that the currents from right to left and from left to right are equal in magnitude and that the total current is

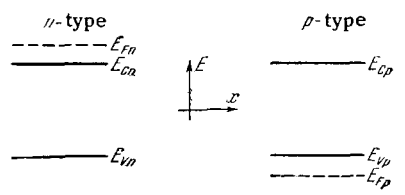

FIG. 3. Position of Fermi level in highly-doped semiconductor.

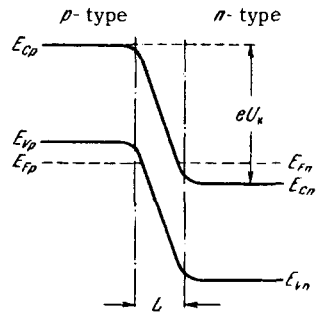

FIG. 4. Energy band scheme (in the absence of bias) in the p-n junction of a tunnel diode.

zero. If a small forward voltage direction is applied to such a p-n junction, the Fermi levels on both sides of the junctions will no longer be on the same horizontal line (Fig. 5a). The number of electrons penetrating from right to left will be larger than the number of electrons passing from left to right. A current will flow through the p-n junction, with a characteristic as shown on the right side of the same figure.

If the bias is increased so that the Fermi level E_{Fp} is on the same horizontal line as the edge of the band E_{cn}, the current increases and reaches a value I_p (Fig. 5b). Further increase in the applied voltage results in a decrease of the current (Fig. 5c), since forbidden levels appear opposite the electrons that could have passed to the left. Finally, if the edges E_{Vp} and E_{cn} are horizontally aligned, the current decreases to zero (Fig. 5d). Further increase in the bias causes ordinary diode current to flow through the p-n junction (Fig. 5e). It is easy to verify that biasing the p-n junction in the opposite direction will produce a strong increase in the current. Thus, the volt-ampere characteristic of the tunnel diode is the sum of two currents—the tunnel current and the ordinary diffusion current. In real devices the tunnel current never decreases to zero. It may reach con-

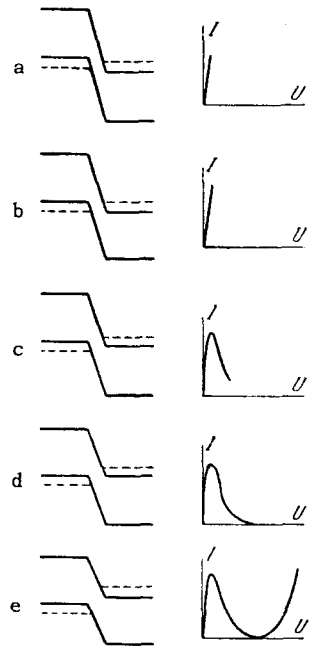

FIG. 5. Energy band scheme and volt-ampere characteristic of a tunnel diode with different values of positive bias.

siderable magnitude, for example one-quarter of I_p in silicon tunnel diodes [7]. The nature of this excess current is still not completely clear and is presently under debate. This question will be discussed in Sec. 5.3.

2. TUNNEL EFFECT IN SEMICONDUCTORS

The tunnel effect is a quantum-mechanical phenomenon. It can be understood by examining Fig. 6, which shows the probability density $|\psi|^2$ of finding a particle in space (for simplicity we chose the one-dimensional case). The same figure shows the thickness of the barrier in an ordinary diode (ac) and a tunnel diode (ab). We see that owing to the narrowness of the latter barrier the probability of finding the particle to the right of this barrier is different from zero.

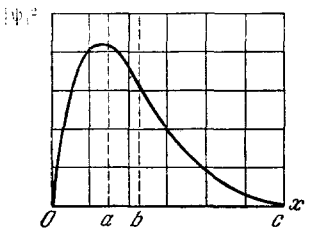

FIG. 6. Probability density of finding an electron in space.

It is clear that the coefficient P characterizing the penetration of the particle through the barrier, say, from left to right will be given by

$$P = \frac{|\psi_{0a}|^2}{|\psi_{bc}|^2}, \qquad (2.1)$$

where ψ_{0a} and ψ_{bc} are the amplitudes of the particle wave functions on the two sides of the barrier, the values of which are determined from the Schrödinger equation.

The first to consider quantitatively the mechanism of the tunnel effect for a semiconductor (but not for a p-n junction) was Zener [8]. The expression for the penetration coefficient, derived in the most correct fashion in [9], has the form

$$P = \exp\left(-\frac{\pi}{2q\hbar\mathscr{E}}\sqrt{2m^*}E_g^{3/2}\right), \qquad (2.2)$$

where m* is the effective mass of the electron and Eg is the width of the forbidden band of the given semiconductor.

This expression was obtained for a barrier of triangular shape. The actual shape of the barrier can hardly have a real meaning, since the de Broglie wavelength of the passing particle should be much shorter than the width of the barrier, i.e., the electric field should stay constant over a distance equal to the de Broglie wavelength. It is easy to show that λ amounts to approximately 30 Å, and that the thickness of the p-n junction, as already indicated, is approximately 100 Å, i.e., the quasi-classical condition is not satisfied. Nonetheless, such an approach enables us to understand the main laws governing the phenomenon.

In order to obtain an expression for the number of electrons penetrating through the barrier per unit volume and per units time, we must multiply (2.2) by the number N of valence electrons per unit volume and by the frequency of collisions between the particles and the barrier. We thus have finally [7]

$$W = N\frac{aq\mathscr{E}}{2\pi\hbar}\exp\left[-\frac{\pi}{2q\hbar\mathscr{E}}\sqrt{2m^*}\cdot E_g^{3/2}\right], \qquad (2.3)$$

where a is the period of the crystal lattice.

It follows therefore that a larger effect will be observed in semiconductors with small values of Eg and m*. For the sake of clarity, Fig. 7 shows the results of calculations with Eq. (2.3) for several semiconductors [10].

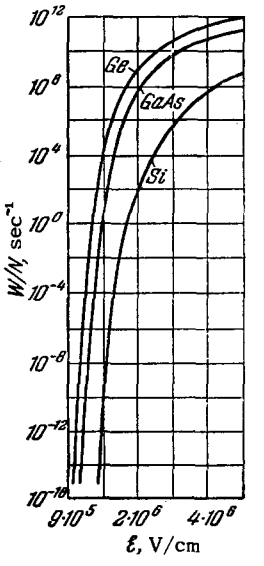

FIG. 7. Probability of electron tunneling as a function of the field intensity in a p-n junction.

The quantum-mechanical theory of the tunnel effect in semiconductors was further developed by Franz and Tewordt [11] and particularly by Keldysh [12,13], who analyzed the three-dimensional problem and obtained for W the expression [12]

$$W = \frac{1}{4}N\frac{a^3q^2\mathscr{E}^2\sqrt{m_{\parallel}}}{\pi\hbar^2\sqrt{E_0}}\exp\left[-\frac{\pi}{2q\hbar\mathscr{E}}\sqrt{m_{\parallel}}E_0^{3/2} + f(\gamma)\right], \qquad (2.4)$$

where $f(\gamma)$ is a factor that accounts for the angular dependence, and $m_{\parallel}$ is some formal quantity defined as

$$m_{\parallel} = \frac{1}{\sum_i (\cos^2\gamma_i)/m_i}, \qquad (2.5)$$

where m_i^{-1} are in turn the principal values of the tensor m_{ik}^{*-1}, while γ are the angles between the field direction and the principal axes of this tensor. Expression (2.4) differs from (2.3) in that the factor preceding the exponential has a different field dependence and the exponential itself has an angle dependence.

The most significant difference, however, lies in the interpretation of the quantities E_0 and $m_{||}$. These quantities assume the respective values E_g and m^* only if the highest state of the valence band and the lowest state of the conduction band correspond to the same value of the quasimomentum (Fig. 8a). In the terminology of Hall et al [14], direct tunneling is obtained. If these states do not coincide, then E_0, identified with the red light absorption boundary of the given semiconductor, will always be larger than E_g.

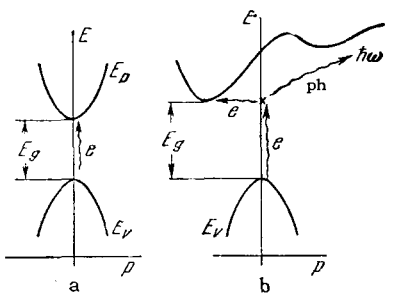

FIG. 8. Tunnel effect in various semiconductors. a) Direct tunneling; b) indirect tunneling.

Consequently the critical fields corresponding to measurable values of W should always be larger than those obtained from (2.3).

Keldysh has shown further that if some interaction is present and changes the quasimomentum of the electron, then the electron can go over into the conduction band even if the maximum of the valence band is shifted relative to the minimum of the conduction band (Fig. 8b). This is called indirect tunneling. The interactions that change the quasimomentum of the electron may be collisions of electrons with one another, with impurity atoms, and with phonons.

In the last-mentioned interaction, the value of W is given by [13]

$$W = A\left(\frac{q\hbar\mathcal{E}}{\sqrt{2m_{||}^*}E_g^{3/2}}\right)^{5/2}\exp\left[f(\gamma)-\frac{4}{3}\frac{\sqrt{2m_{||}^*}}{q\hbar\mathcal{E}}(E_g-\hbar\omega)^{3/2}\right]$$

$$\times\left\{\bar{N}(T)+[1+\bar{N}(T)]\exp\left(-\frac{4}{3}\frac{\sqrt{2m_{||}^*E_g}}{q\hbar\mathcal{E}}\hbar\omega\right)\right\}, \qquad (2.6)$$

where ω is the phonon frequency and $N(\bar{T}) = [\exp(\hbar\omega/kT) - 1]^{-1}$. Unlike in (2.4), here $m_{||}^*$ is the reduced effective mass of the electron and hole.

3. QUANTITATIVE STUDY OF THE TUNNEL EFFECT IN p-n JUNCTIONS

On the basis of qualitative considerations, which pertain to the description of Fig. 5, Esaki [5] proposed that the tunnel current $I_{C\to V}$ from the conduction band of the n-region of the crystal into the valence band of the p-region is proportional to the following: the probability of the tunnel effect, the probability $f_C(E)$ of finding the electron in the conduction band on the right of the junction (Fig. 5) at an energy level E, the probability $[1 - f_V(E)]$ of this level not being occupied in the valence band on the left of the junction, and the densities $\rho_C(E)$ and $\rho_V(E)$ of the energy states in the bands:

$$I_{c\to v} = A_{c\to v}\int_{E_c}^{E_v} W_{c\to v} f_c(E)[1-f_v(E)]\varrho_c(E)\varrho_v(E)\,dE. \qquad (3.1)$$

Analogously, for tunneling from the valence band on the left into the conduction band on the right

$$I_{v\to c} = A_{v\to c}\int_{E_c}^{E_v} W_{v\to c} f_v(E)[1-f_c(E)]\varrho_c(E)\varrho_v(E)\,dE. \qquad (3.2)$$

Esaki considers the difference of these currents to be the total current. He assumes that the bands are parabolic and that the probabilities $W_{C\to V}$ and $W_{V\to C}$ are equal

$$I = B\int_{E_c}^{E_v} W[f_c(E)-f_v(E)]\sqrt{(E-E_c)(E_v-E)}\,dE. \qquad (3.3)$$

This expression, based on general physical considerations only, was given by Esaki without proof, and can therefore not be regarded as a quantitative expression for the volt-ampere characteristic of the tunnel diode. Nonetheless, its simplicity and clarity do cast light on the physical nature of the tunnel diode.

In the particular case when we assume that the p and n regions of the crystals are degenerate to an equal degree, and in such a way that the Fermi level is inside the bands at a distance 2kT from E_V and E_C, respectively, Eq. (3.3) goes over into [15]

$$I = -A\frac{(E_v-E_c)^2\left[1-\exp\left(\frac{qU}{kT}\right)\right]}{[m+n]e^{\alpha/2}\left[1+\exp\left(\frac{qU}{kT}\right)\right]}, \qquad (3.4)$$

where

$$m = \exp\left(-\frac{E_{Fc}-E_c}{kT}\right),\quad n = \exp\left(-\frac{E_v-E_{Fv}}{kT}\right),\quad \alpha = \frac{E_v-E_c}{kT}.$$

A more serious quantitative formulation of Esaki's ideas is the paper by Ivanchik [16], who calculated by the Thomas-Fermi method the variation of the potential in a p-n junction whose both sides are degenerate semiconductors. Knowing the variation of the potential, Ivanchik calculated the tunnel current and obtained a rather cumbersome expression, which is not given here. It is sufficient to state that Ivanchik's formula which he gave in the form of an integral, can be represented by an eighth-degree polynomial; this shows clearly how difficult it is to compare this expression with the experimental results. Qualitatively, on the other hand, it is in satisfactory agreement with the dependence of the tunnel current on the bias (Fig. 9).

Ivanchik's solution covers only direct tunneling.

The quantum mechanical problem of determining the tunnel current through a p-n junction was considered by Price and Radcliffe [17]. They obtained a solution for direct tunneling in the form of an integral

$$I = \frac{2\pi q m_r^* P\mathcal{E}}{ah^3}\int [f_c(E)-f_v(E)]\,g(E_1)\,g(E_2)\,dE, \qquad (3.5)$$

where m_r^* is the reduced effective mass, namely

$$m_r^* = \frac{m_v^* m_c^*}{m_v^*+m_c^*}. \qquad (3.6)$$

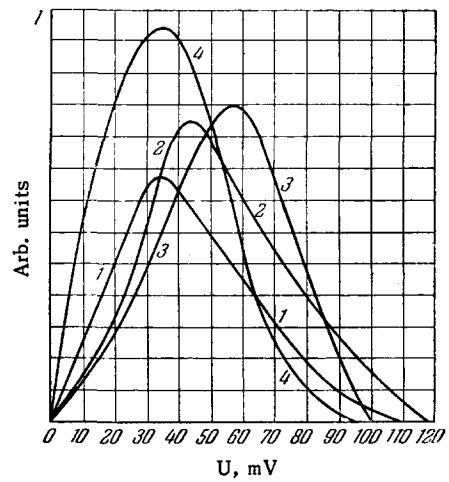

FIG. 9. Calculated dependence of the current on the voltage in the p-n junction of the tunnel diode. 1 – After Esaki; 2 – after Price; 3 – after Ivanchik; 4 – after Kane.

The factors g(E) vanish outside the region where the bands overlap, and are equal to unity inside this region.

For this region we can integrate (3.5) from zero to $E_V - qU$, where U is the applied bias:

$$I = \frac{2\pi q m_r^* kT E_g P}{a\hbar^3 L}\left\{kT\left[\ln\left(1+e^{\frac{E_v-E_F-qU}{kT}}\right)+\ln\left(1+e^{\frac{-E_F-qU}{kT}}\right)-\ln\left(1+e^{\frac{E_v-E_F-2qU}{kT}}\right)-\ln\left(1+e^{-\frac{E_F}{kT}}\right)\right]\right\}. \quad (3.7)$$

If (1.1) and (2.2), which define L and P, are substituted in (3.7), the formula for the volt-ampere characteristic becomes too obscure to an experimenter.

Comparing the calculations made with the formulas of Esaki, Ivanchik, and Price and plotted in Fig. 9, we see that all show a qualitative similarity to the experimentally observed volt-ampere characteristics of the tunnel diode. Unfortunately, the lack of a clear-cut idea of the numerical values contained in the final formulas makes the quantitative comparison of these formulas with experiment difficult.

In our opinion, more fruitful approaches are those of Kane [18] and Bonch-Bruevich [19]. Kane proposed that the electric field in the p-n junction of the tunnel diode is homogeneous. In this case the expression for the tunnel current density has at absolute zero the form

$$j = \frac{qm^*}{18\pi^3\hbar^2}\exp\left\{-\frac{\pi (m^*)^{1/2} E_g^{3/2}}{2\sqrt{2}\hbar\mathcal{E}}\right\}\int\exp\left(-\frac{2E_\perp}{\bar{E}_\perp}\right)dE\,dE_\perp, \quad (3.8)$$

where

$$E_\perp = \frac{\hbar^2(K_y^2+K_z^2)}{2m^*}, \quad (3.9)$$

$$\bar{E}_\perp = \frac{\sqrt{2}\hbar\mathcal{E}}{\pi (m^*)^{1/2} E_g^{3/2}}. \quad (3.10)$$

The factor preceding the integral is the coefficient of electron penetration due to the tunnel effect. In Kane's theory this factor is considered to be constant.

The form of the volt-ampere characteristic therefore depends on the course of the function D(u), which is Kane's notation for the integral in (3.8).

Analytically the function D(u) is represented by Kane in the form of a large number of formulas which describe piece by piece the entire volt-ampere characteristic of the tunnel diode. When Kane's formulas are plotted, all the pieces join continuously into a single smooth curve D(u). Figure 9 shows by way of an example the function D(u) for the case

$$E_{cn} - E_{Fn} = E_{vp} - E_{Fv}$$

in direct tunneling.

Kane also analyzed indirect tunneling, and the analysis can be found in [18].

Bonch-Bruevich [19] considers not the entire volt-ampere characteristic, but only its extremal points. Thus, the voltage corresponding to the maximum tunnel current is given by

$$U_p = \frac{1}{6q}\{2(E_{vp}-E_{Fp})-(E_{Fn}-E_{cn}) + \sqrt{[2(E_{vp}-E_{Fp})+(E_{Fn}-E_{cn})]^2+4(E_{vp}-E_{Fp})(E_{Fn}-E_{cn})}\}. \quad (3.11)$$

At present this approach seems to be the most suitable from the experimenter's point of view. It yields simple expressions that are easy to use, and the quantities in these expressions have a clear cut physical meaning and definite numerical values. It must be noted that the positions of the Fermi levels may be difficult to determine because of formation of impurity bands which merge with the allowed bands in the case of large impurity concentrations.

4. PHYSICAL PRINCIPLES UNDERLYING THE MANUFACTURE OF TUNNEL DIODES

4.1. Semiconductors employed. The qualitative description of the operating principles of the tunnel diode leads to the conclusion that the semiconductors employed in it must be degenerate. Consequently, the Fermi levels of the semiconductors on both sides of the p-n junctions must be not in the forbidden but in the conduction band (for the n type) and in the valence band (for the p type). We can thus determine the minimum carrier concentration at which the electron gas is degenerate. At room temperatures all the impurities in such semiconductors are ionized [20,21], and we can therefore find the degree of doping necessary to realize the tunnel effect.

To be specific, let us take an electron-type germanium crystal. The Fermi level is determined from the curve shown in Fig. 10 [22]. Since the start of the degeneracy is taken to be [23] the condition $E_{Fc} = E_c$, it follows from Fig. 10 that $n_d/N_c = 0.8$. Recognizing that

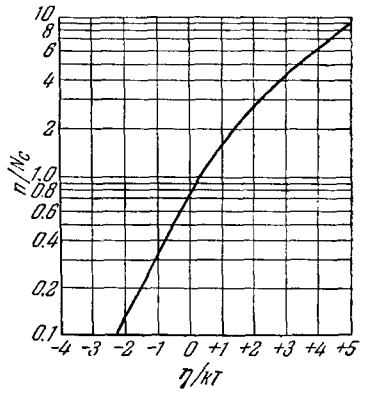

FIG. 10. Dependence of the position of the Fermi level (degree of degeneracy) on the free-electron concentration.

$$N_d = n_d\left(1+2\exp\frac{E_d}{kT}\right) \qquad (4.1)$$

(where E_d is the ionization energy of the donor impurity atoms at large concentrations) we get $N_d = 3n_d$, since $E_d = 0$ [24]. Using the well known expression [2] for N_c, we obtain*

$$N_d = 6\cdot 10^{19}\left(\frac{m^*}{m_0}\right)^{3/2}\frac{T}{300}\,[\mathrm{cm}^{-3}], \qquad (4.2)$$

where T is the temperature in °K. The last column of Table I shows the values of the minimum donor density necessary to obtain a p-n junction with tunnel effect in various semiconductors.

Table I

Semiconductor	$\frac{m^*}{m_0}$	E_g, eV	$n_{eq.}$, cm^{-3}	$(N_d)_{min}$, cm^{-3}
Ge	0.15	0.65	$4\cdot10^{19}$	$3.5\cdot10^{18}$
Si	0.27	1.1	$1.4\cdot10^{20}$	$8.5\cdot10^{18}$
GaAs	0.06	1.35	$7\cdot10^{19}$	$9.0\cdot10^{17}$
InSb	0.04	0.18	$6\cdot10^{17}$	$4.8\cdot10^{17}$

Actually, as will be shown below, the required degree of doping is one order of magnitude higher.

The production of such highly doped semiconductors is a far from simple matter. High degree of doping is possible only with certain impurities with large values of limiting solubility. For germanium and silicon this problem has been investigated in sufficient detail (Figs. 11, 12) [24,25], and for semiconductors that are compounds of group III and V elements of the periodic table (GaAs, InSb, GaSb, etc), the solubility limits have not yet been established for the impurities. There is merely some information [26,27,28] which allows us to state that the most suitable are the n-type systems GaAs + Sn and InSb + Te or the p-type systems GaAs + Zn, GaAs + Cd, and InSb + Cd.

At the same time, the tunnel effect was observed also in gallium arsenide doped with sulfur and selenium [26], although in these cases there is the danger of these elements combining chemically with GaAs.

*The value of N_c used in this case is based on the assumption of Boltzmann statistics and is therefore quite approximate.

High doping methods are considered in [21,29] for germanium and silicon and in [26] for gallium arsenide.

The properties of the crystal determine the density of the tunnel current not in terms of the value of the carrier (impurity) concentration, but also in terms of such parameters as the width of the forbidden band and the effective mass, since these quantities are contained in the expression for tunneling probability (2.2). It is obvious that materials with low effective masses and with low values of E_g need fewer impurities to attain the same current density (3.3). Comparative data on m* and E_g for various materials [30] are listed in Table I.*

The fourth column of this table shows the values of $n_{eq} = n_d n_a/(n_d + n_a)$ needed to attain a tunnel-current density of 10^4 A/cm^2. [31]

4.2. Formation of the p-n junction. The use of a highly doped semiconductor as the initial crystal for the production of a tunnel-effect p-n junction is a necessary but not sufficient condition.

As already indicated, the impurity density on the other side of the p-n junction should also lead to the degeneracy of the electron gas. High density must be combined with a sharp (stepwise) fall-off in the junction region. Only then will the junction be thin and the strong field necessary for the tunnel effect produced in it.

From among the known methods of obtaining p-n junctions such as diffusion, alloying, drawing, etc, the best is the alloying method. It consists of alloying a metal dot, containing an acceptor impurity [32] into an electronic crystal, or a dot containing a donor impurity into a crystal with hole conductivity.

A high impurity concentration is attained in the alloy region primarily by choosing the proper metal. It follows from Figs. 11 and 12 that the most suitable for alloying with electronic germanium are Ga and Al,

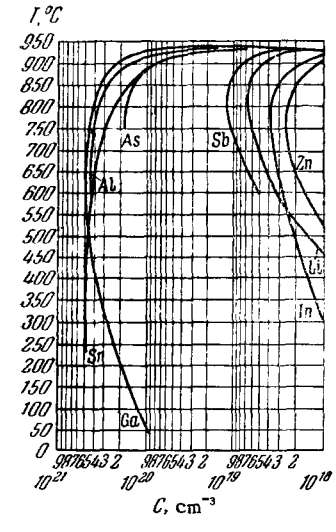

FIG. 11. Temperature dependence of the solubility of certain impurities in solid germanium.

*The considerations advanced are only qualitative, since it is not clear what effective mass is significant in the tunnel effect.

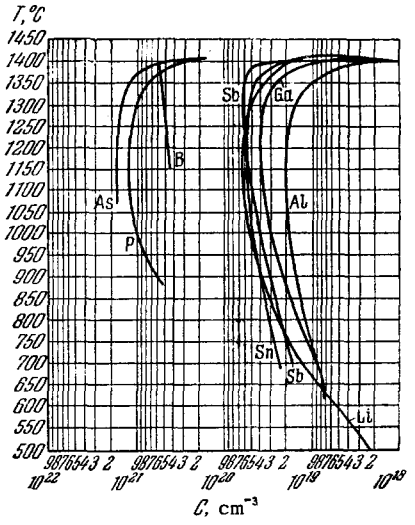

FIG. 12. Temperature dependence of the solubility of certain impurities in solid silicon.

for n-type germanium the most suitable is As, and for silicon— Al, B, and As.

In practice not pure metals but alloys such as InGa [31], SnGa [33], or in the case of silicon AlB [7], are used for alloying.

Naturally, if alloys are used for the alloy junctions, the concentration of the main element that ensures reversal of the sign of the conductivity in the region will be lower than the values given by the solidus curves of Figs. 11 and 12. A rough estimate of the impurity concentration in the alloy region can be obtained in this case by means of the curves of Figs. 13 and 14, which were calculated in [34] and [35] assuming ideal solutions formed upon alloying of binary alloys in germanium. The value of A, the abscissa in these plots, determines the depth of alloying l from the relation

$$l = A\frac{G}{S}, \qquad (4.3)$$

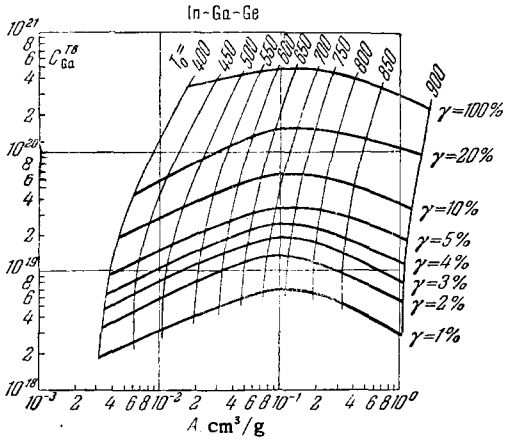

FIG. 13. Distribution of the concentration of gallium in the alloy region of a germanium p-n junction with In-Ga alloy dot.

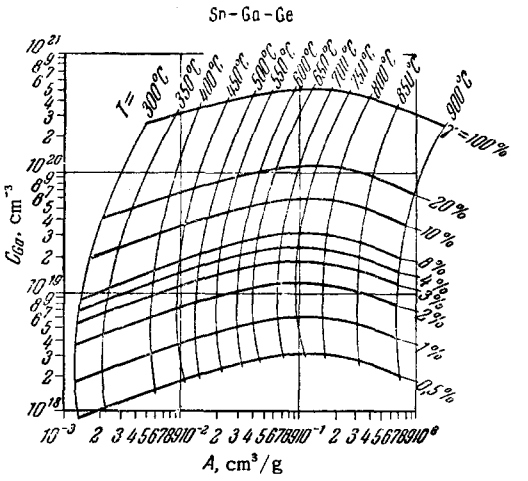

FIG. 14. Distribution of concentration of gallium in the alloy region of a germanium p-n junction with SnGa alloy dot.

where G is the weight of the metal dot and S —area occupied by this dot during alloying.

The quantity γ denotes the percentage content of Ga in the alloys InGa and SnGa.

The curves of Figs. 11 and 12 and the plots of Figs. 13 and 14 are valid only in the case of slow equilibrium cooling of the molten zone in the crystal. This never occurs in practice. To the contrary, it has been shown in many investigations [34,36] that a greater total effect is obtained by abrupt (rapid) cooling. This can be readily explained: in the case of abrupt cooling the content of the impurities in the recrystallized region is greatly increased. An estimate of the concentration is in this case almost impossible, although a knowledge of this concentration is most important, since it is contained in n_{eq}.

The quality of a p-n junction with tunnel effect is greatly influenced by the fusion time, since long heat treatment permits the impurities to diffuse over distances sufficient to smear out the boundaries of the p-n junction appreciably [10].

With the junction having a thickness 100—150 Å, diffusion over a distance ΔL amounting to 10 Å is apparently still tolerable. Then, if the diffusion of the impurities in the semiconductor obeys the Fick law [37]

$$D\frac{\partial^2 N}{\partial x^2} = \frac{\partial N}{\partial t}, \qquad (4.4)$$

where D is the diffusion coefficient, then the maximum permissible heat-treatment time t_{max} at the melting temperature can be estimated from the relation

$$\frac{Dt_{max}}{(\Delta L)^2} \leqslant 1. \qquad (4.5)$$

Figure 15 shows the values of t_{max} as functions of the melting temperature for several semiconductor plus alloy impurity systems. If the correct technology of the alloying process is adhered to, a homogeneous

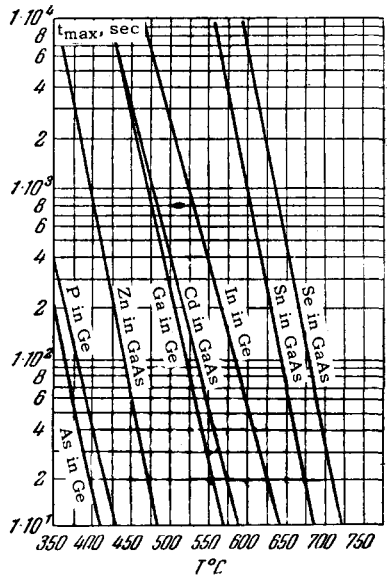

FIG. 15. Permissible heat-treatment time in alloy-junction preparation.

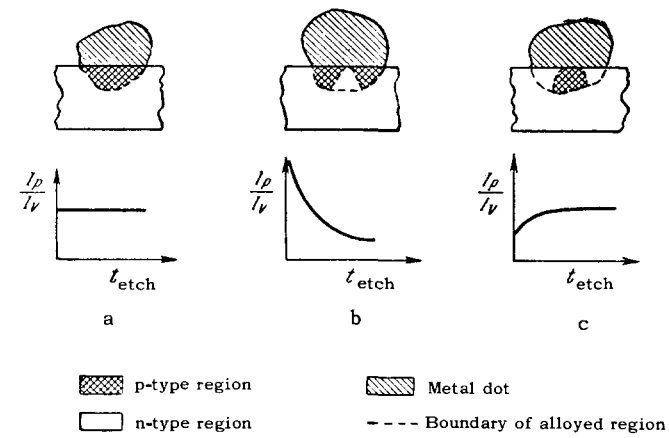

FIG. 16. Types of alloy-region structures in p-n junctions.

region is obtained in the crystal (Fig. 16a). Low-quality processes lead to two types of structural defects (Figs. 16b and c). In the last two cases the current through the crystal will have a non-tunneling component due to the existence of regions comprising a metal and an n-type semiconductor in parallel with the p-n junction. The fraction of this current relative to the total current can be changed by etching the junction.

Accordingly, one can expect also changes in the junction quality, which is best characterized by the current ratio I_p/I_V, as shown in Fig. 16. Such changes have actually been observed in experiments [10,66].

5. PARAMETERS CHARACTERIZING THE TUNNEL DIODE

5.1. Equivalent circuit. To determine the parameters characterizing the tunnel diodes, we turn to the equivalent circuit proposed by Sommers [31] (Fig. 17a).

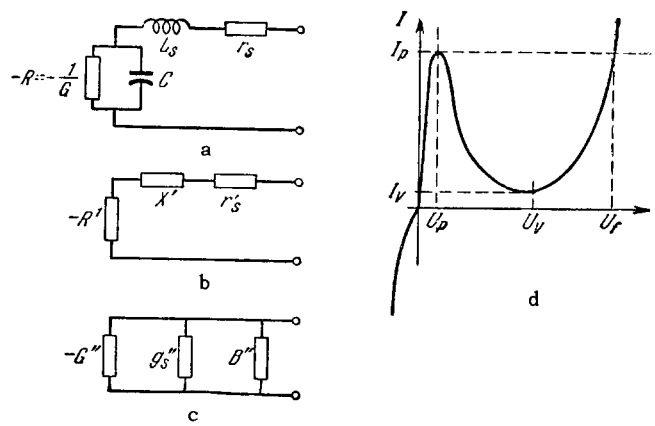

FIG. 17. Equivalent circuit and volt-ampere characteristic of the tunnel diode.

Here C —capacitance of p-n junction, r_S —loss (bulk) resistance, L_S —diode inductance, −R —negative resistance determined by the slope of the decreasing portion of the volt-ampere characteristic of the tunnel diode (Fig. 17d).

The tunnel diode has an "N"-shaped volt-ampere characteristic, a distinguishing feature of which is that the current in the diode has a single-valued functional dependence on the applied voltage and that a point with infinite resistance exists on the boundary between the negative and positive branches (unlike diodes with S-shaped characteristics, where the voltage is a single valued function of the current, and the resistance on this boundary passes through zero). The currents and voltages characterizing the diode are designated on Fig. 17d.

The properties of tunnel diodes as circuit elements are adequately described by the foregoing parameters. [31,38,39] For many applications, however, it is necessary to know the exact form of the volt-ampere characteristic and the voltage dependence of the differential resistance −R and of the capacitance C.

The mixed series-parallel equivalent circuit of the tunnel diode is frequently unsuitable for calculations. All-series and all-parallel circuits are shown in Figs. 17b and c; the transformed equivalent parameters of these circuits are accordingly: [40]

a) for the series circuit

$$R' = \frac{R}{1+\omega^2 C^2 R^2}, \tag{5.1}$$

$$x' = \omega\left(L_s - \frac{cR^2}{1+\omega^2 C^2 R^2}\right); \tag{5.2}$$

b) for the parallel circuit

$$G'' = \frac{R'}{(r_s - R')^2 + x'^2}, \tag{5.3}$$

$$g''_s = \frac{r_s}{(r_s - R')^2 + x'^2}, \tag{5.4}$$

$$B'' = -\frac{x'}{(r_s - R')^2 + x'^2}. \tag{5.5}$$

Sometimes, however, it is convenient to leave R and

G alone, and to refer to their terminals all the other diode and circuit parameters.

5.2. Peak current. The concentration dependence of the peak current I_p is derived from the concentration dependence of the tunnelling probability.

Indeed, calculation[41] yields for I_p an equation similar to (2.2):

$$I_p = I_0 S \exp\left[-BE_g\left(\frac{\varepsilon m^*}{n_{eq}}\right)^{1/2}\right], \qquad (5.6)$$

where I_0 —proportionality coefficient (in current-density units), B —a constant dependent on the choice of units, and S —area of the p-n junction.

Expression (5.6) shows that for a given semiconductor the maximum current depends most strongly on the equivalent doping. This conclusion is satisfactorily confirmed by experiment, as can be seen from Fig. 18 which we plotted using the data of [42,43]. It is interesting to note that doping the crystal with various impurities leads to different values of I_p (even if the concentration is the same). Thus, Furukawa[43], noticed an increase of the p-n junction tunnel current in germanium doped with arsenic, as compared with junctions in antimony-doped germanium. This agrees also with the values of the electron mobility in these cases[43,44]. There is still no satisfactory explanation for this fact.

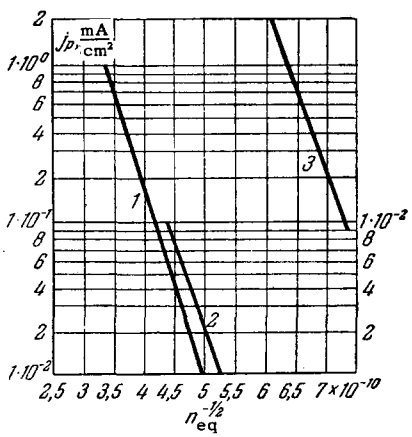

FIG. 18. Dependence of the current density in the tunnel diode on the equivalent doping. 1 – Initial crystal Ge + As; 2 – initial crystal Ge + Sb, right-hand current scale; 3 – initial crystal Ge + P.

Since (5.6) contains the quantity n_{eq}, it is natural to expect I_p, and consequently also I_p/I_V, to depend strongly on the impurity concentration in the alloy region of the crystal. It is impossible, however, to trace this dependence by direct means, since the concentration cannot be determined. This dependence was observed indirectly in many investigations[21,34,42,45].

In Fig. 19, taken from [45], the unknown concentration is implicitly contained in the values of the U_p ordinates. This diagram serves to illustrate the distribution curves. It shows that diodes with different U_p, and likewise with different I_p/I_V, can be obtained for the same concentration n_d. This diagram is very useful for the choice of the initial germanium crystal needed to obtain a specified I_p/I_V. Unfortunately,

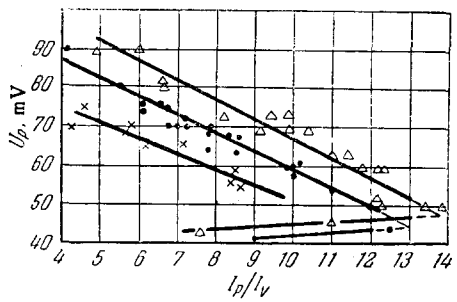

FIG. 19. Curves illustrating the distribution of the tunnel diodes by values of I_p/I_V. Δ, •, and × – 3.5, 2.5, and 1.8×10^{19} donor atoms/cm³, respectively.

there are no similar diagrams for other materials.

It is easy to deduce from (3.3) that the tunnel current should vary with the p-n junction temperature.

The current variation is due to the temperature dependence of the Fermi levels in both the p and the n regions of the crystal and to the temperature dependence of the forbidden band width. These two dependences are shown schematically in Fig. 20 for n-type germanium. The Fermi level comes down when a certain rather low temperature T_c is reached, and the degree of degeneracy is accordingly reduced. A decrease in the forbidden band width, however, increases the degree of degeneracy. The temperature interval connected with the degeneracy will be smaller in a material with relatively low impurity content, and consequently the temperature variation of E_F is the predominant factor. In highly doped specimens, however, the second mechanism of temperature action can no longer be discarded.

In the first experiments on tunnel diodes[5,21,46] the semiconductors were relatively weakly doped ($n_d \leq 1 \times 10^{19}$ cm^{-3}), and a decrease in the tunnel current with increasing temperature was therefore observed. Later investigations[42,47] disclosed also the second type of temperature dependence.

The temperature dependence of the tunnel current becomes most clearly manifest in investigations of the value of I_p. Figure 21 shows the relative temperature variation of I_p for a germanium tunnel diode, as given by Cady[45]. The voltage U_p is again used as the parameter.

In those cases when indirect tunneling occurs, the tunneling probability should increase with temperature, as follows from (2.6). Therefore, even in the case of

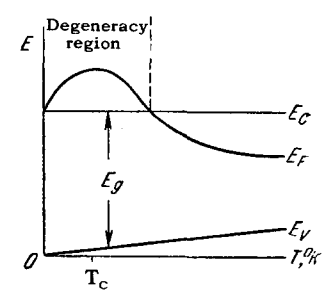

FIG. 20. Temperature variations of the Fermi level and of the forbidden band width in a degenerate semiconductor.

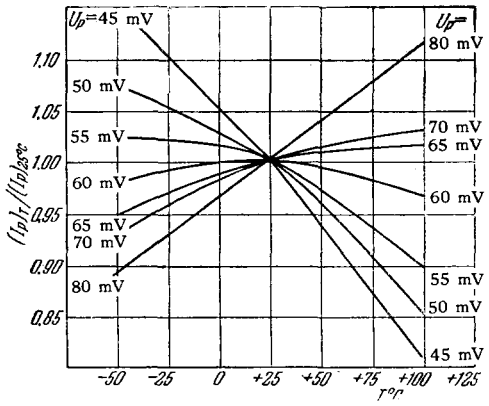

FIG. 21. Temperature variations of the current I_p.

crystals with relatively low impurity content, the current may increase with temperature rather than decrease. This occurs, for example, in antimony-doped germanium, where an increase of current with increasing temperature is observed even at 5×10^{18} cm^{-3} of antimony [47]. This is in remarkable agreement with the experimentally observed electron-phonon interaction during the course of tunneling in similar p-n junctions [48].

So far there are no reports of a systematic investigation of the temperature dependence in tunnel diodes based on compounds of the A^{III}-B^{V} type or on silicon. However, the tunneling in A^{III}-B^{V} compounds is expected from the band structure [49] of the semiconductors to be direct, and the dependence to be thus qualitatively similar to that shown in Fig. 21. In the case of silicon diodes, where the tunneling is indirect [48], the current increases with increasing temperature at all concentrations.

The working-temperature interval of tunnel diodes is determined essentially by the width of the forbidden band and therefore varies with the type of semiconductor (see Table II).

5.3. Excess current. The simplified treatment of the band structure in the p-n junction (Fig. 5e) predicts a decrease of the tunnel current to zero. In real diodes, however, this is not observed and the current at the voltage U_V may be large.

In one of the first papers on tunnel diodes [50] Esaki advanced the idea that this excess current is due to "tunneling." He believed that the deep levels in the forbidden band, with which the excess current is connected, are due to dislocations. Later experiments have established that I_V is independent of the pressure [52] and of the temperature [53], and that the donors and acceptors have a rather weak concentration dependence [54]. All these experimental facts favored tunneling as the cause of excess current. By now, on the basis of additional data [53,55], a specific model has been proposed [54] to explain the tunneling through the deep levels. This model is shown schematically in Fig. 22, which gives several possible tunneling mechanisms:

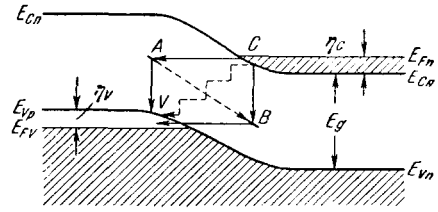

FIG. 22. Explanation of the occurrence of excess current in a tunnel diode.

1. The electron moves from the conduction band on the right to the level A, from which it then goes into the valence band.

2. The electron moves from C to the deep level B from which it penetrates into the valence band.

3. The electron enters the valence band along the path CABV in the presence of an impurity band.

4. A multi-stage transition of the electron from C through the local levels, with radiation of energy.

A tunneling mechanism with energy radiated in the form of photons or phonons was examined theoretically by Kane [56], who showed that such transitions cannot occur in practice. No photons were observed in any of the experiments, although the procedures employed had a large sensitivity margin (more than one order of magnitude [54]).

The most probable path is CBV. Tunneling from B to V is none other than the field ionization theoretically considered by Franz [57], who showed that the transition probability will have the same form as for intraband transitions if E_g is replaced by the impurity ionization energy. With such a transition, Chynoweth et al [54] derived the following expression for the excess current

$$I_{ex} = \ln A + \ln D_{ex} - \frac{\alpha_{ex} L^{1/2}}{2} [E_g - qU_{ex} + q(\eta_V + \eta_C)], \qquad (5.7)$$

where D_{ex} is the density of the states (of the levels B) in the forbidden band; U_{ex} —bias voltage (volts) where the excess current predominates; η_V and η_C —positions of the Fermi levels (in volts) in the corresponding crystal zones, and

$$\alpha_{ex} = \theta \left[\frac{8}{3} \frac{(m_{\parallel}^{*})^{1/2}}{qh} \right], \qquad (5.8)$$

where θ is a factor close to unity.

Expression (5.7) is easily verified by experiment, since D_{ex}, L, E_g, and U_{ex} can be varied independently and their effect on the behavior of I_{ex} investigated.

A detailed study [54] has demonstrated quite convincingly the validity of (5.7) for silicon diodes, and consequently demonstrated the tunneling nature of the excess current. Finally, the reliability of the model considered was conclusively confirmed once an additional negative-conductivity portion was observed in the excess-current region at nitrogen temperatures [58]. Moreover, special doping of the germanium crystal with gold [59] has led to the observation of two maxima on the excess-current portion, corresponding to the

two 0.35 and 0.54 eV levels of gold below the bottom of the conduction band. In exactly the same manner, Chynoweth [59] recorded five maxima and related them with the phosphorus vacancy levels.

An interesting investigation of the impedance Z of a GaAs tunnel diode at microwave frequencies was undertaken by Esaki [59]. He observed that Z oscillates as a function of the frequency in the excess-current region. The period of these oscillations was approximately 10^{-8} sec. The physical interpretation of this time is unclear, but assuming the equivalent circuit in this case to be a delay element, Esaki believed that what he measured was the time that the electron remained at the B level (Fig. 22).

The determination of the impurity levels in the forbidden zone from the appearance of additional maxima on the volt-ampere characteristic of a tunnel diode has come to be known as "tunnel spectroscopy." It is quite possible that similar experiments will develop into a usable method for the general study of semiconductors. Nor is the possibility of technical utilization of the additional negative-conductivity regions on the tunnel-diode characteristic excluded.

5.4. Characteristic voltages. The nature of the tunnel currents I_p and I_V is associated with voltages U_p and U_V corresponding to these currents. In first approximation, the characteristic voltages are independent of the methods used to manufacture the diodes. A more rigorous analysis shows that both U_p and U_V are determined by the value of n_{eq}, i.e., the impurity concentrations on both sides of the p-n junction.

We do not know the explicit analytic dependence of U_p and U_V on the degree of degeneracy in both parts of the crystal. It has been shown experimentally that U_p increases somewhat with increasing impurity concentration in the initial crystal [34,42].

The change in U_p due to the voltage drop on r_s must be taken into account in such investigations.

The voltage U_V also increases with increasing impurity concentration in the n and p regions of the crystal [42]. This agrees with the ideas developed in [19].

The character of the temperature variation of U_p and U_V is entirely a consequence of the temperature dependence of I_p and I_V. The value of U_p decreases very slightly (for germanium—by 5—10 mV) with increasing temperature [5,7,21,42,47]. U_V is practically independent of the temperature until the increase in the diffusion component of the current begins to have an appreciable effect (the right-hand rising branch of the volt-ampere characteristic of the diode). In the latter case, U_V also decreases with increasing temperature.

Typical values U_p and U_V for tunnel diodes made of different semiconductors are listed in Table II. The same table shows the current ratio I_p/I_V and the parameter U_f, which correlates with the width of the forbidden band.

5.5. Negative resistance. The negative resistance is not a resistance in the ordinary sense of the word. But since the differential equations that describe the behavior of the system with positive and negative damping are analogous, the concept "negative resistance" is universally employed.

Negative resistance (or negative damping) leads to an increase of energy in the system and is made possible by the presence of an energy source (in the case of the tunnel diode, this source is the bias battery which fixes the operating point of the diode on the negative-slope portion of the volt-ampere characteristic). The tunnel-diode resistance seen by this battery is always positive. But for certain values of the bias corresponding to the negative slope the tunnel diode is a negative differential resistance. For brevity we call this simply the negative resistance.

The foregoing detailed analysis of the current parameters of the diode predetermines the nature of the variation of the negative resistance $-R$.

For a given material, the minimum negative resistance varies uniformly in inverse proportion to the tunneling probability (2.2). Taking (1.1) into account, we can see that the larger the free-carrier (i.e., impurity) concentration on both sides of the p-n junction, the smaller the minimum negative resistance. In practice the lower limit of R is determined by the maximum solubility of the impurities in the semiconductor crystal and by the possibility of obtaining the most differentiated distribution of the concentration near the p-n junction. We note that R is inversely proportional to the area of the junction.

5.6. Capacitance of the p-n junction. The capacitance of the p-n junction of the tunnel diode is given by the well known expression [22]

$$C = 1{,}05S\left[\frac{\varepsilon q n_d n_a}{8\pi (U_k - U)(n_d + n_a)}\right]^{1/2}, \qquad (5.9)$$

Table II

Semiconductor	I_p/I_V	U_p, mV	U_V, mV	U_f, mV	T_{max}, °C	RC, sec	I_p/C, mA/pF
Ge	10—15	40—70	270—350	450	250	$0.5\cdot10^{-9}$	0.3—1
Si	3—4	80—100	400—500	700	400	$0.2\cdot10^{-8}$	<0.5
GaAs	40—70	90—120	450—600	1000	600	$0.1\cdot10^{-9}$	10—15
GaSb	15—20	30—50	200—250	450	300	$0.1\cdot10^{-9}$	—
InSb	7—10			200	25	$0.5\cdot10^{-11}$	

where S is the area of the junction; the remaining symbols are the same as in (1.1).

This dependence was repeatedly checked [61-63] over a wide range of voltages and was shown to be valid up to bias values at which the ordinary diode current comes into play [63]. The absolute value of the capacitance changes by only 20—30 in this voltage range.

It is characteristic that the ratio $k = I_p/C$ is a factor independent of the area, and is essentially a function of n_{eq} only.

This has enabled Cady [45] to plot the curves of Fig. 23 for germanium tunnel diodes. It follows from this figure, in particular, that if n-type germanium is used it is difficult in practice to produce tunnel diodes with a performance factor $I_p/C > 1$, but this can be done by using other materials. Thus, if GaAs is used, the factor k amounts to about 10 mA/pF. [64]

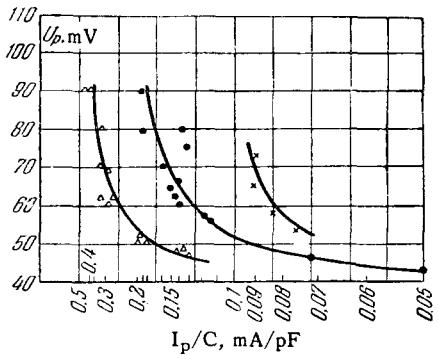

FIG. 23. Curves illustrating the distribution of tunnel diodes by values of I_p/C. Δ, •, and × – 3.5×10^{19}, 2.5×10^{19}, and 1.8×10^{19} donor atoms/cm³.

Typical values of I_p/C for diodes of different types are shown in Table II.

5.7. Time constant. A second factor independent of the junction area and dependent only on the equivalent doping is the diode time constant RC, which increases almost exponentially with the free-carrier concentration [31,63]. Figure 24 shows this dependence for diodes made of Ge and GaAs.

The values of RC normally observed in diodes made of different semiconductors are listed in Table II.

5.8. Loss resistance. In view of the use of low-resistivity semiconductors in tunnel diodes, the loss (bulk) resistance r_s is small. Thus, for germanium tunnel diodes r_s measured with direct current is on the order of several ohms or a fraction of an ohm.

The values of r_s in tunnel diodes cannot be estimated with the formulas for ordinary diodes, because the semiconductors on both sides of the tunnel junction have approximately equal resistivities. In ordinary diodes the main semiconductor layer always has high resistivity and determines principally the loss resistance. Moreover, calculation shows [65] that if the tunnel diode is prepared by alloying an acceptor impurity in n-type germanium, almost all the loss resistance

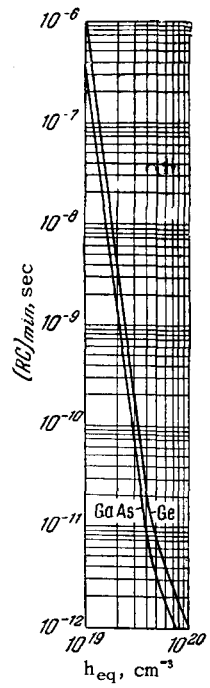

FIG. 24. Dependence of the tunnel diode time constant on the equivalent doping.

is concentrated in the alloy region. In this case the value of r_s should be inversely proportional to the p-n junction area. This indeed was observed in the experiments [66].

The quantity r_s/R is a performance factor of the tunnel diode, dependent on n_{eq}, as is the time constant RC.

There is still no detailed mathematical analysis of r_s as a function of the geometry and nature of the semiconductors employed.

5.9. Limiting and resonant frequencies. The highest frequency at which a tunnel diode can generate is determined by equating to zero the active part of the input impedance (Fig. 17b)

$$Z_{in} = \frac{-R}{1+\omega^2C^2R^2} + r_s + j\omega\left[L_s - \frac{CR^2}{1+\omega^2C^2R^2}\right] \qquad (5.10)$$

and can be written

$$\frac{R}{1+\omega^2C^2R^2} \geqslant r_s. \qquad (5.11)$$

This yields the well known equation for the limiting oscillation frequency [31]

$$\omega_{lim} = \frac{\sqrt{\frac{R}{r_s}-1}}{RC}. \qquad (5.12)$$

The natural resonant frequency of the tunnel diode, determined by equating to zero the imaginary part of the input impedance is

$$\omega_0^2 = \frac{1}{L_sC}\left(1 - \frac{L_s}{CR^2}\right). \qquad (5.13)$$

Tunnel diode microwave devices operating above the natural resonance ω_0 are difficult to build. It is therefore desirable to shift the natural resonant frequency of the diode beyond the frequency region in which its

impedance is active. This is expressed by the inequality

$$\omega_0 \geqslant \omega_{\text{lim}}. \tag{5.14}$$

6. TUNNEL DIODE CONSTRUCTION

In addition to the ordinary requirements imposed on the construction of semiconductor diodes, a tunnel diode must also have low inductance L_S. In accordance with (5.14), the inductance L_S should be lower than

$$L_{\max} = CR_s^2. \tag{6.1}$$

For this equation to be satisfied, the typical parameters of germanium tunnel diodes ($RC \sim 10^{-9}$ sec, $r_S \sim 1$ ohm) would call for an inductance less than 10^{-9} H. Ordinary diodes, in which the contact with the p-n junction is by means of thin wires, have inductances not less than 3×10^{-9} H. Thin wires are therefore not used for the electric leads of tunnel diodes.

The electric contact is produced either with a membrane [10] (Fig. 25a), by clamping to a bulky electrode [67] (Fig. 25,b) or with a strip [68] (Fig. 25c).

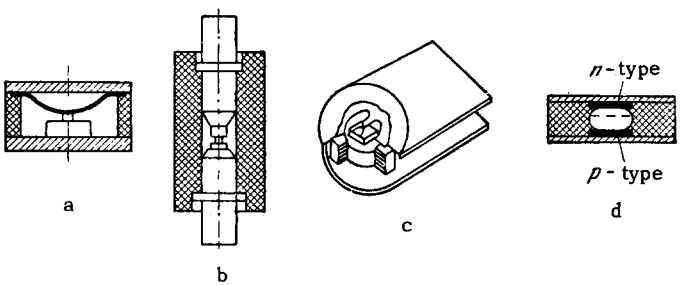

FIG. 25. Types of tunnel-diode constructions.

In the latter construction L_S is 4.0×10^{-10} H.

The best low-inductance electric contact is produced by a flat plate directly soldered to the alloyed metal dot (Fig. 25d). In this case the tunnel-diode inductance drops to a value determined by the geometry of the crystal itself with an alloyed metal dot.

7. OPERATING CONDITIONS AND STABILITY OF TUNNEL-DIODE CIRCUITS

Tunnel diodes can be used in amplifiers, oscillators, converters, and switches. The necessary relations between the diode parameters and the circuit parameters in these devices are determined by stability analysis. These relations can be determined by examining the very simple circuit in which a tunnel diode is connected to the bias source through a resistor r and an inductance L (the values of L and r include the loss (bulk) resistance r_S and the diode inductance L_S) (Fig. 26). The steady-state equation is

$$U + ri_d = E, \tag{7.1}$$

the graphic solution of which plotted together with the

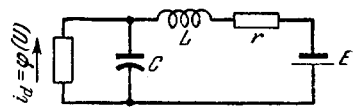

FIG. 26. Equivalent circuit of a tunnel diode connected to a bias source; $r = r_S + r_1 + r_l$, $L = L_S + L_1$ (the subscripts s, 1, and l refer to the diode, the resonant circuit, and to the load, respectively).

volt-ampere characteristic of the diode on Fig. 27, determines the equilibrium state of this circuit.

The behavior of the simplified circuit (Fig. 26) on the negative branch of the characteristic is described by a second order differential equation [69]

$$\left.\begin{aligned} i &= C\frac{dU}{dt} + \varphi(U), \\ L\frac{di}{dt} &+ ri + U = E; \end{aligned}\right\} \tag{7.2}$$

$$\frac{d^2U}{dt^2} + \left(\frac{r}{L} - \frac{G}{C}\right)\frac{dU}{dt} + \frac{U}{LC} + \frac{r}{LC}\varphi(U) + \frac{1}{C}\frac{dC}{dU}\left(\frac{dU}{dt}\right)^2 = E\frac{1}{LC}. \tag{7.3}$$

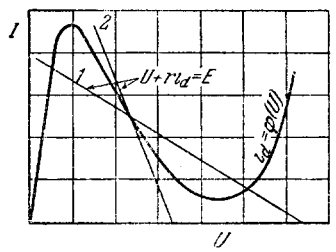

FIG. 27. Graphic determination of the equilibrium states of the circuit in Fig. 26. 1) $r > 1/G$, equilibrium at point a always unstable; 2) $r < 1/G$, stability of the equilibrium at the point a depends on the fulfillment of the additional condition $r > LG/C$.

According to Lyapunov, (7.3) can be linearized to determine the stability of the equilibrium points against small disturbances.

The linearized equation is [60]

$$\frac{d^2U}{dt^2} + \left(\frac{r}{L} - \frac{G}{C}\right)\frac{dU}{dt} + \frac{1 - rG}{LC}U = 0, \tag{7.4}$$

or, changing to dimensionless parameters

$$\ddot{U} + (p - s)\dot{U} + (1 - ps)U = 0, \tag{7.5}$$

where

$$\omega_0 = \frac{1}{\sqrt{LC}}, \quad p = \omega_0 rC, \quad s = \omega_0 LG,$$

and the dots denote differentiation with respect to $t_1 = \omega_0 t$.

The stability of the equilibrium states is checked by an analysis of the roots of the characteristic equations corresponding to the differential equations (7.4) or (7.5). The equilibrium is stable only if the real parts of both roots of the characteristic equations are negative. (The same stability conditions follow from the requirement that the impedance have no zeroes in the right half of the complex-frequency plane [61,70].)

If

$$ps > 1 \quad \left(r > \frac{1}{G}\right) \tag{7.6}$$

the roots of the equation are real and have opposite signs; the equation describes a system with repulsion force, characterized by an unstable equilibrium of the saddle type, regardless of the relations between the remaining parameters of this system.

In this mode, the tunnel diode can be used for switching. When

$$ps < 1 \quad \left(r < \frac{1}{G}\right) \tag{7.7}$$

the load line crosses the volt-ampere characteristic at one point only and the dc stability conditions are fulfilled.

Complex roots occur for the following inequality

$$p + s < 2, \tag{7.8}$$

when the integral curves on the phase plane have the form of one spiral wrapped in another. The asymptotic point of all the curves is a focus; the system can support periodic oscillations. The stability of the focus is determined by the Lyapunov method. When

$$p > s \quad \left(r > \frac{LG}{C}\right) \tag{7.9}$$

the equilibrium is stable. The region of the stable focus contained between the abscissa axis and the lines p = s and p + s = 2 is characterized by damped sinusoidal oscillations. They can exist if the following conditions are fulfilled

$$p > s, \quad p + s < 2. \tag{7.10}$$

The mode corresponding to this region is used in amplifiers: when

$$p < s \quad \left(r < \frac{LG}{C}\right), \quad p + s < 2 \tag{7.11}$$

the focus is unstable. Periodic oscillations can be generated in the mode corresponding to the region contained between the ordinate axis and the lines p = s and p + s = 2.

In the region between the hyperbola ps = 1 and line p + s = 2, the equilibrium state is a node, and the process is aperiodic. Stability is determined as in the case when p > s. The possible operating modes of the tunnel diode are illustrated in Fig. 28.[60]

We note that to satisfy the stability conditions, Eq. (7.9) must be satisfied at all frequencies at which the tunnel diode is active.

In the case of frequency-dependent $L(\omega)$ and $r(\omega)$, expression (7.9) has the form

$$r(\omega) > \frac{L(\omega)G}{C}. \tag{7.12}$$

In the general case, the system can be more complicated than that considered in Fig. 26.

The description of such systems leads to differential equations of order higher than the second. A direct calculation of the roots of such equations is difficult. The stability conditions can be determined in this case from the coefficients of the characteristic equation by means of the Routh-Hurwitz criterion (see, for example,[71]).

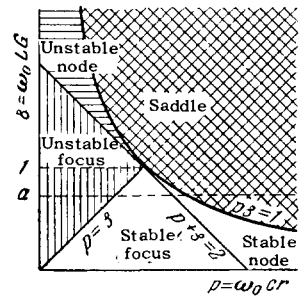

FIG. 28. Stability diagram of tunnel diode.

In some cases, for example when the resonant frequencies of the individual parts of the circuit differ greatly from each other, it is possible to subdivide the complicated circuit into simpler ones. This was done by Hines[73] to analyze a circuit with blocking capacitance (see Sec. 9.3).

8. MEASUREMENT OF TUNNEL-DIODE PARAMETERS

8.1. Volt-ampere characteristic. An investigation of the volt-ampere characteristic is a far from simple problem, in view of the presence of a negative resistance in the tunnel diode and in view of the need for satisfying the stability conditions at all frequencies from zero to the limiting frequency.

The characteristic is determined either point by point or with an oscilloscope. To determine the current, a resistance r_1 is connected in series with the diode and the voltage drop on it measured. This circuit is shunted by resistance r_l, to which a voltage is applied (dc if the characteristic is plotted point by point, or pulsating if an oscilloscope is used)[61]. In this simplest circuit the dc and ac resistances are equal.

To obtain an undistorted negative branch, the characteristic should be plotted in a mode characterized by a stable focus or node. For this purpose it is necessary (but not sufficient) to ensure that the maximum diode and circuit inductance not exceed the value

$$L_{\max} < \frac{C}{G^2}, \tag{8.1}$$

which follows directly from the condition s < 1 (Fig. 28). Then the dependence of the mode on the resistance $r = r_1 + r_S + r_l$ is described by the diagram of Fig. 29, which represents the intersection of the stability diagram by the line s = a (with a < 1) (assuming that the resistance of the power source connected to r_l is large compared with r_l). The stable region occurs if conditions (7.7) and (7.9) are satisfied:

$$\frac{1}{G} > r > \frac{LG}{C}. \tag{8.2}$$

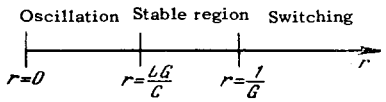

FIG. 29. Dependence of the operating modes of the circuit of Fig. 26 on the active resistance r (with s < 1).

If the circuit oscillates, r is increased until (7.9) is satisfied and the oscillation stops. If this calls for such large values of r that the dc stability conditions are not satisfied, the negative branch of the characteristic cannot be obtained without distortion. (This distortion is connected with the detection of the resultant oscillations and their superposition on the volt-ampere characteristic.)

In a circuit with blocking capacitor (Fig. 30), the capacitor C_b and the resistor r_l are separated in space, so that temperature measurements can be made and the inductance of the resistor r_l can be made insignificant. [74]

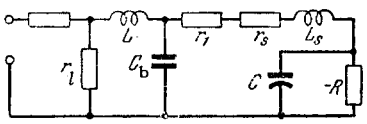

FIG. 30. Diagram for plotting the volt-ampere characteristic.

The stable region can be expanded by connecting a shunting resistor directly to the terminals of the tunnel diode. By adjusting the value of this resistor, the total resistance of the shunted tunnel diode can even be made positive. Electric neutralization of the fraction of the current flowing through the shunting resistor r is by means of a bridge circuit [75]. This bridge, the diagonal of which is connected to the vertical plates of the oscilloscope tube, is balanced with the tunnel diode disconnected. When the scope is connected, the beam deflection represents the current flowing through the diode.

We note that the volt-ampere characteristics can be readily observed with the circuit of Fig. 36. In this circuit [73] the dc and ac stability conditions are different and are easier to obtain (the load r_l on the high frequency circuit does not affect the dc stability).

8.2. Measurement of the negative resistance. Strictly speaking, the differential negative resistance −R cannot be measured directly. The measured quantity, $-R + r_s$, is subject to considerable error if $r_s \ll R$. In the opposite case, r_s should be measured separately and accounted for in the determination of the negative resistance. This resistance can be determined by drawing a tangent to each point of the volt-ampere characteristic (to avoid errors, the stability conditions (8.2) must furthermore be satisfied. [61]

The approximate value of the negative resistance at the maximum-slope points can be determined by measuring the extremal values of the currents I_p and I_V and the corresponding voltages U_p and U_V, using the formula

$$R_{\min} \approx K \frac{U_V - U_p}{I_p - I_V}, \qquad (8.3)$$

where K = 0.4 for germanium tunnel diodes.

To avoid large errors, the negative resistance should be measured with an impedance bridge (Fig. 31) at a frequency low enough to satisfy the conditions [76]

$$\omega^2 R^2 C^2 \ll 1,$$
$$\omega^2 (L_s - CR^2)^2 < (r_s - R)^2, \qquad (8.4)$$

which follow directly from (5.1)—(5.5). (The lead inductance is disregarded here.) To prevent distortion of the results by the averaging, the amplitude of the tunnel diode voltage must not exceed several millivolts.

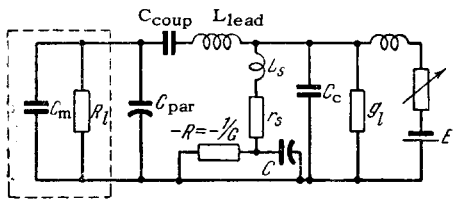

FIG. 31. Circuit for the measurement of negative resistance and capacitance of a tunnel diode with the aid of an impedance bridge.

A convenient method of determining negative resistance is to measure the current flowing through the diode and the voltage on the diode in response to a small ac signal.

8.3. Measurement of the junction capacitance. It is possible to measure the junction capacitance at any point of the volt-ampere characteristic by using an impedance bridge [61]. The apparent capacitance at the input of the instrument is

$$C_{app} = C + C_{par} - L_s G^2. \qquad (8.5)$$

This approximate equation holds true if conditions (8.4) are satisfied.

In a real measuring circuit, it is necessary to take into account the loading of the diode by the bias source conductance g_l, the inductance L_{lead} of the lead between the tunnel diode and the bridge input and the capacitance of the tunnel diode cartridge C_c. Account of these supplementary factors leads to the expression [77]

$$C_{app} = C + C_c - L_{lead}(g_l + G)^2 - L_s G^2. \qquad (8.6)$$

In a high frequency diode, the intrinsic inductance is small and L_{lead} may be the main source of error, particularly in measurements in the negative-resistance region, where a large value of g_l must be used to insure dc stability.

The accuracy of the method increases appreciably if the measurements are confined to the vicinity of the valley on the positive portion of the characteristic, where a small value of g_l can be tolerated. [77]

The cartridge capacitance C_c can be measured separately and accounted for in subsequent calculations.

The capacitance of the tunnel diode, at voltages corresponding to the peak current or the valley current,

can be simply measured by a method based on the divider principle. In this method the signal is applied to a series circuit consisting of the active resistance r and the tunnel diode biased to one of the extremal points of the volt-ampere characteristic (Fig. 32).

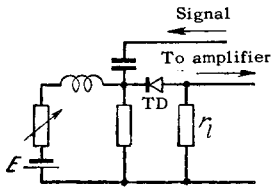

FIG. 32. "Divider" circuit for the measurement of tunnel-diode capacitance.

If the signal is small and a suitable frequency chosen so that the circuit inductance and the loss resistance can be neglected, the diode admittance at these points is determined only by the junction capacitance. Measurement of the capacitance at a voltage corresponding to the peak current imposes more stringent limitations on the signal magnitude and on the stability than the voltage corresponding to the valley.

This method can be used to measure the capacitance of the tunnel diode in the vicinity of the negative branch of the characteristic. For this purpose the negative conductance of the diode is shunted with a resistance of a value that makes the resultant conductance equal to zero.

The capacitance at the extremal points of the characteristics can be also measured by a resonance method, which reduces to a determination of the resonant frequency of the tunnel-diode capacitance and a known inductance L (much larger than the diode inductance L_S) (Fig. 33) at voltages just preceding the onset of oscillation. A suitable frequency for the measurement of the capacitance of a high-frequency tunnel diode is on the order one megacycle.

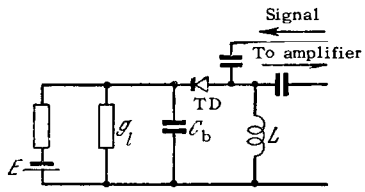

FIG. 33. Circuit for the measurement of tunnel-diode capacitance by the resonant method.

8.4. Measurement of loss resistance (r_S) of a tunnel diode. In dc and low-frequency measurements, only an approximate value of r_S can be measured, since the skin-effect losses do not come into play. The resistance r_S shifts the peak voltage on the volt-ampere characteristic towards the higher values [61]. This effect is appreciable in large-current diodes. For quantitative measurements this method can hardly be used, since the peak voltage, being a function of the doping in the absence of r_S, is not known accurately. The tunnel-current region is not suitable for the measurement of r_S, since the conductance is not determined by ohmic losses in this region.

It is customary to measure r_S on the negative branch of the volt-ampere characteristic at currents that are 100—200 times larger than the peak current, when the shifting of the bands in the p-n junction does not limit these currents. The measurements can be made either by determining the slope of the volt-ampere characteristic or by employing a small ac signal [61]. Measurements on the forward rising branch of the characteristic (beyond the current valley) are considered to be crude, in view of the distortion of the results by secondary effects such as admittance modulation. According to the data of [76], these effects are negligibly small and measurements on the forward branch give results that do not differ from measurement results on the negative branch of the volt-ampere characteristic.

8.5. Measurements of series inductance of the tunnel diode. These measurements are among the most difficult, for the smallness of this inductance makes bridge methods little effective. [61]

A possible method of measuring L_S is to determine the circuit parameters in a mode close to free oscillation. [78] The gist of the method is to place the diode in a special short circuited measuring line of variable length which is set to oscillate. The length of the line is decreased until the oscillation frequency approaches the limiting diode generation frequency in the particular circuit. Near the point where the generation stops the circuit oscillations are sinusoidal with low amplitude, and their frequency is measured with the aid of a sensitive wave meter.

In such a circuit the diode inductance is given by

$$L_s = \frac{1}{\left(\omega_{0\max}^2 + \frac{G^2}{C^2}\right)C} - L_{\min}, \qquad (8.7)$$

where $\omega_{0\max}$ —maximum diode generation frequency in the test unit and $L_{\min}$ —corresponding minimum value of the loop inductance.

It is possible to measure in the same circuit the loss resistance of the diode, provided all the other circuit losses can be accounted for.

A convenient method of measuring very small inductances ($L_S < 10^{-9}$ H) is based on measuring the frequency of resonance between the inductance and the cartridge capacitance. This resonant frequency, which usually lies in the microwave band, is determined from the phase difference between the standing-wave minimum of a short circuited line and the line containing the tunnel diode. If the effect of the junction capacitance on the resonant frequency cannot be neglected, this capacitance must be taken into account in the inductance calculation. The same method can be used to determine the loss resistance r_S at microwave frequencies, by measuring the standing wave ratio at the resonant frequency and by suitable recalculation.

9. TUNNEL DIODE GENERATORS

9.1. General relations. The generation of periodic oscillations corresponds on the diagram of Fig. 28 to an unstable focus region. The solution of the nonlinear differential equation (7.3) determines the frequency and amplitude of the tunnel-diode generator with equivalent circuit of Fig. 26.

This equation, however, cannot be solved in general form. The corresponding linearized equation (7.4) yields the oscillation frequency provided the smallness of the steady-state amplitude allows us to neglect the effect of the nonlinearity of the volt-ampere characteristic on the frequency. In the opposite case it determines the initial frequency.

Of course, the linearized equation yields no information regarding the amplitude, for any deviation in a linear system should eventually increase without limits. The initial frequency is determined by the second term of the radicand of the characteristic equation

$$\lambda = -\frac{1}{2}\left(\frac{r}{L}-\frac{G}{C}\right) \pm \sqrt{\frac{1}{4}\left(\frac{r}{L}-\frac{G}{C}\right)^2 - \frac{1-rG}{LC}}. \quad (9.1)$$

The steady state frequency ω_S differs from the initial frequency

$$\omega = \left(\frac{1-rG}{LC}\right)^{1/2} \quad (9.2)$$

because of the nonlinearity of the diode, and can be determined in terms of the effective steady-state parameters R_{de} and C_{de} [79]:

$$\omega_s = \left(\frac{1-\frac{r}{R_{de}}}{LC_{de}}\right)^{1/2}, \quad (9.3)$$

$$L = rR_{de}C_{de}. \quad (9.4)$$

The exact values of R_{de} and C_{de} can be determined only by numerical or graphical methods [80]. However, if the oscillation amplitude is small, the changes in R and C are small and their effective values can be replaced by the average values $\overline{R}$ and $\overline{C}$. [79] Then

$$\omega_s \approx \left(\frac{1-\frac{r}{\overline{R}}}{L\overline{C}}\right)^{1/2} = \left(\frac{1}{L\overline{C}} - \frac{r^2}{L^2}\right)^{1/2}, \quad (9.5)$$

$$L \approx r\overline{R}\,\overline{C}. \quad (9.6)$$

The smaller the oscillation amplitude, the less the average values differ from the initial parameters at the operating point. Equation (9.2) can therefore be regarded, in any case, as the solution of the first approximation for large-amplitude oscillations.

The total solution of the nonlinear differential equation can be obtained with desired accuracy (including the determination of the frequency, the amplitude, the harmonic content, and the wave form of the oscillations) by either numerical or graphical methods. [10,79] The use of computers greatly accelerates the solution. Thus, a digital computer, together with approximation of the tunnel-diode volt-ampere characteristic by the relation

$$I = A_1U + A_2U^2 + A_3Ue^{-2/3}\left(\frac{U}{A_4}\right)^{3/2} + A_5(e^{A_6V} - 1) \quad (9.7)$$

and suitable programming, yielded results that agreed qualitatively with experiment [81].

Steady state occurs at the operating frequency ω of the circuit of Fig. 26 when conditions (5.12) and (5.13) are satisfied simultaneously. In these expressions it is necessary to add to the resistance r_S the load resistance r_l and the circuit loss resistance r_1, while L_S can be regarded as the limiting value of the joint inductance L of the diode and the circuit.

Thus, the generation condition at frequency ω has the form

$$\omega = \frac{\sqrt{\frac{R}{r_s+r_l+r_1}-1}}{RC} = \sqrt{\frac{1}{LC}\left(1-\frac{L}{CR^2}\right)}. \quad (9.8)$$

For a diode characterized by parameters C, R, and L_S, it is necessary to choose the corresponding value of r_l at each frequency.

9.2. Generator circuits. The systems usually employed to generate microwave oscillations with the aid of tunnel diodes contain line elements with distributed parameters. The literature reports only one type of generator with lumped parameters with the tunnel diode forming an integral part of the generator [33].

Historically, the first device for the generation of microwave oscillations was the circuit proposed by Sommers [31]. In this generator the low-inductance tunnel diode was connected to the terminals of an open strip line without additional connecting leads (Fig. 34). This made it possible to shunt the diode by means of a small resistance at the node of the high-frequency voltage and obtain dc bias without disturbing the high-frequency operating conditions.

To prevent oscillations in the circuit made up of the diode, the line segment, and the shunting resistance, the resonant frequency of this circuit must be shifted above the limiting frequency of the tunnel diode. This circuit is an example showing that in view of the broadband nature of the negative resistance, one of the main difficulties encountered in the design of tunnel-diode generators, is the elimination of all the undesirable frequencies.

It can be shown that the input admittance of the line (Fig. 34) has at the junction with the tunnel diode, disregarding r_S and L_S, the form

$$Y_{in} = \frac{1}{W}\operatorname{th}(\alpha + j\beta)l \approx \frac{1}{W}\left(\frac{\alpha l}{\cos^2\frac{2\pi l}{\lambda}} + j\operatorname{tg}\frac{2\pi l}{\lambda}\right) \text{ for } \alpha \ll 1, \quad (9.9)^*$$

where W —characteristic resistance of the line, α — damping factor, and β —phase-shift constant.

When $\lambda/2 > l > \lambda/4$, this admittance is inductive. The length of the line is determined from the condition that the imaginary part of the input admittance must be cancelled by the capacitance of the tunnel diode

*th = tanh, tg = tan.

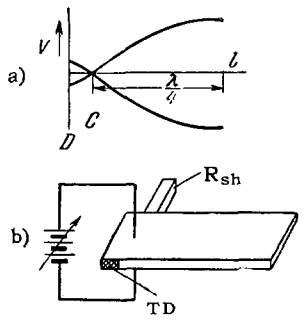

FIG. 34. Schematic representation of the first tunnel-diode generator described by Sommers. a) Distribution of voltage along the line; b) TD – tunnel diode; R_{sh} – resistor.

$$j\omega C + \frac{j}{W}\operatorname{tg}\frac{2\pi l}{\lambda} = 0. \qquad (9.10)$$

In a somewhat modified generator circuit [79] the tunnel diode is connected to an open line shunted on one of its ends by a stabilizing impedance (Fig. 35). The distance l_1 is chosen to make the conductance of the line at the junction to the diode equal to the modulus of its negative admittance. The length of the open loop l_2 is chosen to cancel out the reactive component. The total admittance of the line (in the TEM mode) at the junction with the diode is given by the equation [79]

$$\frac{Y_c}{Y_0} = \frac{\frac{W}{Z_R}}{\left(\frac{W}{Z_R}\right)^2 \sin^2\beta l_1 + \cos^2\beta l_1} + j\left[\operatorname{tg}\beta l_2 + \frac{\left\{1-\left(\frac{W}{Z_R}\right)^2\right\}\sin\beta l_1\cos\beta l_1}{\left(\frac{W}{Z_R}\right)^2\sin^2\beta l_1 + \cos^2\beta l_1}\right], \qquad (9.11)$$

where Y_0 is the characteristic admittance of the line and Z_R is the stabilizing impedance.

Similar relations can be established also in the case when a ring-type strip resonator is used [79]. Plots of (9.11) and of the corresponding equation for the ring resonator show that tunnel-diode generators can be constructed with parameters that differ appreciably.[81]

A Rieke diagram can be drawn for a tunnel-diode generator to show the variation of the frequency and power with the load resistance. The feasibility of electrical tuning of a tunnel diode generator follows from Eq. (9.8) and was experimentally confirmed by many workers.

Thus, in [82] a frequency tuning range of 12% was obtained by varying the position of the operating point on the volt-ampere characteristic.

9.3. Methods of applying bias to the tunnel diode. The use of a blocking capacitor [73] to apply the bias improves the low-frequency stability and facilitates construction of tunnel-diode generators. In particular, it makes it possible to employ short-circuited loops in place of open ones in the circuit described above. The advantages resulting from the use of a blocking capacitor to stabilize a negative-resistance circuit can be understood by analyzing the equivalent circuit of the generator, shown in Fig. 36. In this figure C_b is the blocking capacitor and r_l is the load resistance, which is inductively coupled to the circuit.

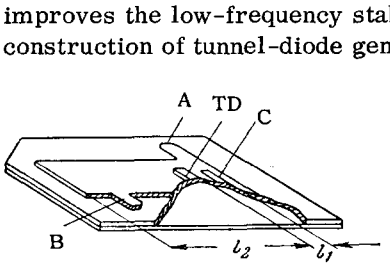

FIG. 35. Schematic representation of generator with stabilizing resistor. A – Input for bias source; B – high frequency output; C – stabilizing impedance; TD – tunnel diode.

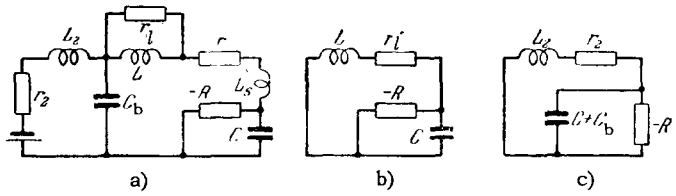

FIG. 36. Circuit with blocking capacitor to apply bias to a tunnel diode. a) Actual circuit; b) high-frequency equivalent circuit; c) low-frequency equivalent circuit.

The stability condition for the equivalent circuit of the generator at low frequency (Fig. 36c) has the form

$$R > r_2 > \frac{L_2}{(C+C_b)R} \quad \text{or} \quad \sqrt{\frac{L_2}{C+C_b}} < R. \qquad (9.12)$$

It is easier to satisfy (9.12) by connecting a large blocking capacitor C_b. It is necessary to take account, however, of the fact that this capacitor is part of the high frequency circuit and must have low losses. If the impedance of the bias source is not low enough to satisfy the dc stability conditions, it should be reduced by means of an additional shunting resistance.

It should be noted in this connection that the use of a voltage divider for a bias source with low output impedance has low power efficiency.

Thus, for example, to obtain a bias of 0.2 V across a one-ohm resistor from a two-volt battery, the power loss required is 0.4 W, which is approximately two orders of magnitude in excess of the power necessary to operate a germanium diode with 50 mA peak current. Thermoelectric converters have low internal impedance and apparently solve the problem of supplying tunnel diodes.

Thermoelectric converters can be fed with direct or alternating current, have negligible output-voltage pulsations, high efficiency, and an output voltage that can be easily regulated. Converters can be made with internal resistance not exceeding 0.02 ohms.[83]

9.4. Problems involved in increasing the power rating of tunnel-diode generators. In first approximation it can be assumed that the volt-ampere characteristic of the tunnel diode is linear in the range from U_p to U_V and that the power delivered by the diode is (when $r_S = 0$)

$$P = \frac{1}{8}(U_p - U_V)(I_p - I_V). \qquad (9.13)$$

The actual volt-ampere characteristic is close to linear only on a certain part of the negative branch. Confining ourselves to this region, which is of impor-

tance for microwave generators, we obtain an expression for the output power at a frequency ω, when $R \gg r_s$, in the form

$$P = \frac{V}{8L} \frac{K}{\omega^2 + \frac{K^2}{V^2}} \left(1 - \frac{\omega^2}{\omega_{\text{lim}}^2}\right), \qquad (9.14)$$

where V is the voltage range in which the volt-ampere characteristic does not deviate from linearity by a specified amount (say 15%), and $K = I_p/C$. (According to [84], typical values for germanium tunnel diodes are $V \approx 0.1$ V and $K = 0.3-1$ mA/pF.) Expression (9.14) enables us to analyze the dependence of the power on the diode and circuit parameters.

The results of such an analysis are presented in graphical form in [84], where it is shown that for each given frequency $\omega = \omega_0$ there is an optimum $K = K_0$ at which the output power of the diode is maximal

$$P_{\max} = \frac{V^3}{16K_0L} = \frac{V^2}{16\omega_0 L}. \qquad (9.15)$$

This value of K_0 is given by the equation $\omega_0 = K_0/V$.

To increase the power is one of the main tasks in the development of tunnel-diode generators. To obtain oscillation power exceeding several milliwatts is a problem involving serious difficulties.

Thus, to obtain a power on the order of a milliwatt from a typical germanium diode it is necessary to employ a diode with a negative resistance of one ohm or less. Merely to stabilize the operating points of such a tunnel diode is a very difficult problem.

With increasing operating frequency, the problem of increasing the power becomes even more serious. It is obvious that the fraction of the power P delivered to the load r_l decreases relative to the fraction of the loss power P_s in the diode and in the circuit;

$$\frac{P}{P_s} = \frac{r_l}{r_s + r_1} = \frac{\frac{R}{1 + R^2\omega^2C^2} - (r_s + r_1)}{r_s}. \qquad (9.16)$$

The impossibility of using diodes with large peak currents is due to the decreased resonant frequency of diodes that have small R and large C.

On the other hand, diodes characterized by very small values of junction capacitance have an intolerably large value of r_s. Therefore the development of diodes with small junction capacitances and loss resistances is of prime significance. This problem can be solved either by using welded p-n junctions [85] or by producing very thin semiconductor layers [84].

The maximum power realized at the present time is several milliwatts at frequencies up to 1 Gc and fractions of a milliwatt at frequencies up to 10 Gc. In many applications this power level is fully adequate; a certain increase in these levels can be expected in the future. Nonetheless, the foregoing difficulties make it necessary to seek other solutions if the output power is to be increased appreciably.

We note that the prospects of further appreciable increase in the power would involve the development of tunnel diodes having long negative branches of the volt-ampere characteristic. The output power can be increased by operating two or more tunnel-diode generators in parallel [81]. Parallel operation is made possible by the frequency locking phenomenon.

Figure 37 shows the frequency dependence of the synchronous gain g_s (ratio of the output power of parallel-connected generators to the power necessary for locking) on the frequency difference Δf between the generated oscillations and the synchronizing signal. This curve shows that synchronization can be produced with a signal that has a power several orders of magnitude less than that of the synchronized generators [81]. In such a parallel generator, the tunnel diodes can be connected to the common circuit by means of hybrid rings or circulators [86].

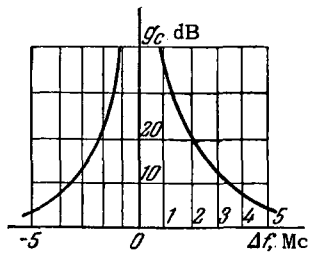

FIG. 37. Dependence of the synchronous gain on the difference between the frequencies of the generated oscillations and the synchronizing signal.

The output power of a circuit with a hybrid ring coupled to three tunnel-diode generators reached 5.6 mW at a 850 Mc, or 90% of the sum of the powers of the three generators [81].

A rather promising method of increasing the power is to use a distributed tunnel diode in which the p-n junction is quite wide. Certain possibilities of using the distributed tunnel diodes were considered theoretically by Hines [73]. The simplest "smooth" distributed diode of appreciable width, in which the wave propagates in the barrier layer between p-n junctions that form a sort of transmission line, is not very suitable for practical applications. The reasons are the low limiting frequency, the very low impedance, and the lack of means for connecting the fixed bias.

Greater possibilities are afforded by a very narrow strip diode (Fig. 38) with a negative resistance located along the central axis of a metallic strip line of considerable width. Such a structure supports a TE mode and the phase velocity is smaller than that of light, in view of the capacitive load of the strip-type tunnel diode.

For such waves, the magnetic and electric fields

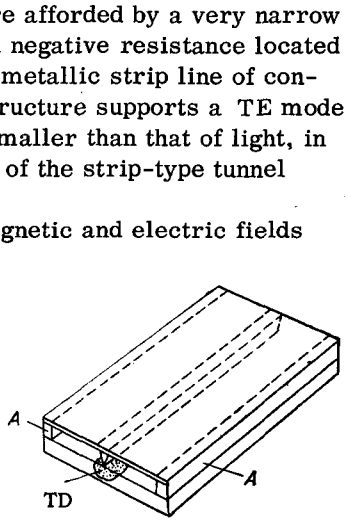

FIG. 38. Connection of "distributed" diode to a strip line. TD – "narrow" strip diode; A – absorbing loads.

in a direction perpendicular to the strip attenuate exponentially and in proportion to the frequency. Therefore two absorbing loads placed parallel to the strip axis and a certain distance from the axis provide a considerable conductance for all low-frequency oscillations (including direct current) without influencing noticeably the high-frequency fields.

In this simple generator, the narrow strip diode wavelength is equal to half the wavelength in the given structure. Such a resonant tunnel strip is placed perpendicular to the cross section of the waveguide and excites oscillations in the latter. Another possible method is to excite a cavity by using two distributed strips on both sides of an inductive slot in this cavity. In the "useful" type of oscillations employed, all the "elementary" diodes of the given strip oscillate in phase while the oscillations in the two tunnel strips are out of phase. It is also possible to use distributed structures with axial symmetry.

9.5. Experimental results. The practical generator circuits differ from one another in the principal mode employed, in the method used to couple the system to the external load, and in the method used to apply the bias to the tunnel diode.

The resonant circuits employed are segments of coaxial and strips lines [31,73,79], waveguides of narrow cross section [87,88], H-shaped waveguides [79], toroidal and conical cavities [87] and ring cavities [79]. In all cases, the design of the oscillating system involves neutralization of the reactance (principally capacitive) of the tunnel diode. The neutralization methods also vary (inductive short-circuited or open-circuited loops, inductance of a toroidal-cavity turn, etc).

The coupling with the load can be capacitive, inductive, or conductive; most frequently the tunnel diode is connected directly in the waveguide [79,87,88]. The customarily used method of applying the bias consists of using a blocking capacitor [73].

Let us examine certain practical constructions of tunnel-diode generators and the experimental results obtained.

An overall view of a coaxial-type generator with blocking capacitor for the bias and with coupling to the cavity through a diaphragm is shown in Fig. 39. Oscillations were produced in such a construction [73] at frequencies up to 8 Gc; the same device can be used as an amplifier. A certain practical inconvenience in such a construction is the difficulty of regulating the coupling to the load and the need for producing an artificial return circuit for the dc bias of the tunnel diode.

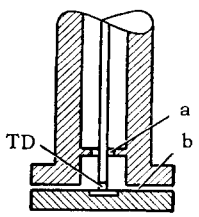

FIG. 39. Coaxial oscillator with blocking capacitor. TD – tunnel diode; a – diaphragm to couple the cavity with the line; b – blocking capacitor.

Cavities and waveguides are used as tank circuits for tunnel-diode oscillators of higher frequencies. [79, 87,88]

Frequencies up to 33.4 Gc were obtained [87] with a welded GaAs tunnel diode, produced directly in the center of a conical cavity. The oscillation was excited at the higher cavity modes. A germanium tunnel diode used in the same cavity at 2.7 Gc oscillated only on the lowest cavity mode. The oscillation spectrum contained harmonics in addition to the fundamental frequency. The bias was applied through a parallel plate blocking capacitor and the energy was drawn through a coupling loop. A frequency of 40 Gc was obtained by using a segment of tapered waveguide as the tank circuit. The reactive component of the diode was neutralized with the aid of an adjustable short-circuited plunger located in the waveguide past the diode. The frequency decreased monotonically with increasing bias. However, large-current tunnel diodes are used to generate oscillations of large amplitude and high harmonic content, the frequency can go through a maximum or a minimum, depending on the bias.

The use of an analogous welded diode of n-type GaAs has made it possible to raise the oscillation frequency to 103 Gc. [88] The diode was produced in a 3.1 × 0.25 mm waveguide. Fundamental oscillations starting with frequencies close to waveguide cut-off up to frequencies more than three times this value were observed in one and in the same system, depending on the diode and on the position of the tuning plunger.

Certain interest attaches to the production of the SS 104 and SS 107 experimental tunnel-diode oscillators announced by RCA. [89,90] The SS 104 oscillator has a 5% electric tuning range relative to a selected frequency in the 800—1400 Mc band. The oscillator draws 30 mA at 0.2 V and has a maximum output power of 0.3 mW. The output power of the SS 107 oscillator exceeds 1 mW at a fixed frequency in the 0.5—2 kMcs band. This oscillator draws 160 mA at 0.4 V.

The highest frequency obtained with a germanium tunnel diode is 108 Gc. [66] In this experiment the diode parameters were $L_S = 0.25 \times 10^{-9}$ H, C = 0.784 pF, R_S = 6 ohms, R = 55 ohms, and the oscillation frequency corresponds to the value given by Eq. (9.8). [84] According to [130], the D4168-D p-type germanium tunnel diode can also generate at frequencies above 10 Gc.

The reported experimental data confirm that in principle there is no upper frequency limit in the tunnel-effect mechanism and justify the bold predictions made in earlier papers that tunnel diodes can be used in the millimeter band.

10. TUNNEL-DIODE AMPLIFIERS

Tunnel-diode amplifiers are of the regenerative type. Amplification corresponds to the region char-

acterized by a stable focus in the diagram in Fig. 28.

10.1. Amplifier with separate input and output ("straight through" circuit). The gain of a tuned amplifier represented by the series or parallel equivalent circuit of Fig. 40, is defined as the ratio of the power delivered to the output load to the maximum power delivered by the signal source.

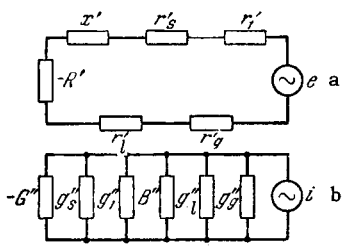

FIG. 40. Equivalent circuits of "straight through" resonant amplifier. a) Series circuit, b) parallel circuit.

The gain expressed in terms of the parameters of both circuits has the form

$$k=\frac{4r'_g r'_l}{(r'_g+r'_s+r'_1+r'_l-R')^2}=\frac{4g''_g g''_l}{(g''_g+g''_s+g''_1+g''_l-G'')^2}. \quad (10.1)$$

In these equations r'_S, r'_1, r'_g, r'_l, g'_S, g''_S, g''_1, g''_g, and g''_l are the loss resistances and conductances of the diode, the tank circuit, the oscillator, and the load in the series and parallel circuits, respectively. In the latter circuit the conductances are referred to the terminals of the signal generator (g''_g) and the load (g''_l).

Taking into account the increase in the effective Q of the tank circuit resulting from the regeneration, and the relationships between the real and transformed diode parameters, we obtain for the product of the voltage gain $\sqrt{k}$ by the amplifier bandwidth Δf

$$\sqrt{k}\Delta f=\frac{1-\frac{r_s+r_1}{R'}}{\pi CR\left(\sqrt{\frac{r'_l}{r'_g}}+\sqrt{\frac{r'_g}{r'_l}}\right)}. \quad (10.2)$$

In the derivation of (10.2) it is assumed that the gain is large. The equation can be rewritten in a form more suitable for practical use

$$\sqrt{k}\Delta f=\left(1-\frac{Q_l}{Q_0}\right)\left(1-\frac{r_s}{R}\right)\times\left(1-\frac{\omega^2}{\omega^2_{\lim}}\right)\frac{1}{\pi CR\left(\sqrt{\frac{r'_l}{r'_g}}+\sqrt{\frac{r'_g}{r'_l}}\right)}, \quad (10.3)$$

where

$$1-\frac{Q_l}{Q_0}=\frac{r'_g+r'_l}{r'_g+r'_1+r'_l} \quad (10.4)$$

is the efficiency of the cavity.

In this form the equation is a particular case of a more complete solution [91], which takes into account the finite nature of the gain and the effect of the distributed inductance of the cavity.

The optimum relation between the load and generator resistances, from the point of view of increasing the product $\sqrt{k}\,\Delta f$, occurs when $r'_g = r'_l$. However, to decrease the noise figure it is best to have $r'_g \gg r'$ (see Sec. 10.6). Neglecting the losses r_S and r_1, we obtain the well known approximate formula [31]

$$\sqrt{k}\Delta f\approx\frac{1}{2\pi RC}. \quad (10.5)$$

Methods for determining the optimum gain-bandwidth product and for the synthesis of amplifiers, not limited to the foregoing system with a single resonant circuit, are given in [92-94]. The gain-bandwidth product can be increased appreciably for similar amplifiers constructed in the form of multi-element filters. The results of the calculations for such an optimal amplifier, for the case of a "straight through" circuit, are shown in Fig. 41 together with a generalized circuit of this amplifier. [92]

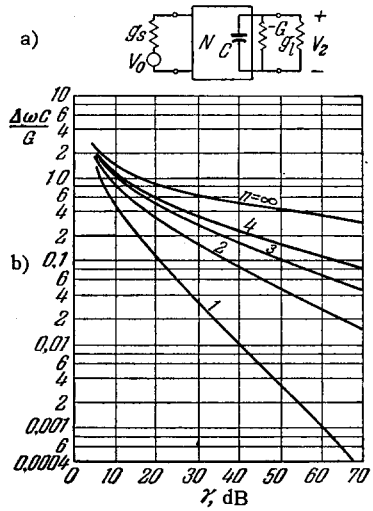

FIG. 41. Dependence of the bandwidth on the gain for an optimal amplifier (b) in a "straight through" circuit (a). n – number of poles of the filter $\Delta\omega = 2\pi\Delta f$.

We note that unlike simple tuned amplifiers, the gain-bandwidth product in optimal amplifiers is no longer invariant. [94]

10.2. Amplifier with circulator in "reflection" circuit. The use of a "reflection" circuit with a circulator offers appreciable advantages from the point of view of expanding the bandwidth, improving the stability, and reducing the noise figure. This connection is consequently most widely used for microwave amplification. In the circuit of Fig. 42, g_{loss} and g_0 are the loss conductance and the conductance of the input line, respectively, transformed to the terminals of $-G$ and B, where B is the susceptance of the amplifier (the primed values G" and $g''_{loss} = g''_S + g''_1$, encountered in the text, correspond to conductances transformed to the terminals of a signal generator with conductance g''_0).

The gain γ is defined as the square of the modulus of the voltage reflection coefficient $|\Gamma|^2$ [2]:

$$|\Gamma|^2=\gamma=\frac{(g_0-g_{loss}+G)^2+B^2}{(g_0+g_{loss}-G)^2+B^2}. \quad (10.6)$$

The gain exceeds unity if G is larger than g_{loss}, and reaches infinity at resonance (B = 0) if $G = g_0$

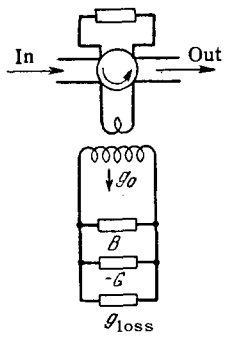

FIG. 42. Amplifier with circulator in a "reflection" circuit.

+ g_{loss}. Oscillation takes place when $G > g_0 + g_{loss}$.

As in other regenerative amplifiers, the gain-bandwidth product remains approximately constant at large gains, when $g_0 \approx G - g_{loss}$. In this case

$$\sqrt{\gamma} = \frac{2(G - g_{loss})}{g_0 - (G - g_{loss})}, \quad (10.7)$$

$$\Delta f = f_0 \frac{g_0 + g_{loss} - G}{\omega C}, \quad (10.8)$$

$$\sqrt{\gamma}\Delta f \approx \frac{G - g_{loss}}{\pi C}. \quad (10.9)$$

If the losses are small, the product $\sqrt{\gamma}\,\Delta f$ is

$$\sqrt{\gamma}\Delta f \approx \frac{1}{\pi RC}. \quad (10.10)$$

Thus, the "reflection" amplifier circuit with circulator has twice the gain-bandwidth product of the "straight through" amplifier circuit.

The results of calculation of an optimum amplifier for a "reflection" circuit with circulator are shown in Fig. 43.[92] A comparison of the curves in this figure with those of Fig. 41 shows the advantages of the "reflection" circuit in the case when complicated resonant circuits are used.

10.3. Balanced tunnel-diode amplifier. Efforts to prevent the noise figure from being adversely affected

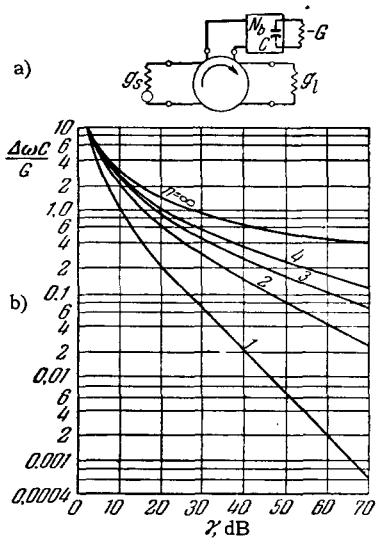

FIG. 43. Dependence of the bandwidth on the gain for an "optimal" amplifier (b) in a "reflection" circuit with circulator (a). n – number of poles of the filter $\Delta\omega = 2\pi\Delta f$.

by load noise entering the amplifier without resorting to ferrite circulators, have led to the development of a balanced amplifier circuit. Such a circuit, first proposed for paramagnetic amplifiers[95], was subsequently realized also for tunnel-diode amplifiers.[96] (A balanced amplifier is sometimes also called "hybrid.") The power reflected from two identical amplifiers located in the arms of a hybrid junction (Fig. 44) is summed in arm 4, and the amplified receiver noise is summed in arm 3. The gain of the balanced circuit when the input to the hybrid junction is matched is given by Eq. (10.6). The relationships between the gain and the bandwidth shown in Fig. 43 for the "reflection" amplifier with circular also hold for the balanced amplifier.[92]

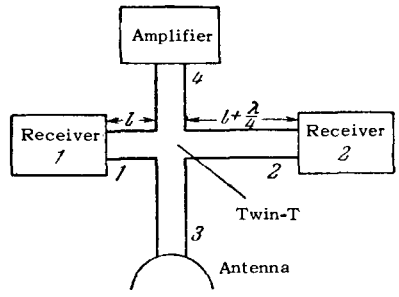

FIG. 44. Schematic diagram of balanced amplifier.

With the input and output impedances matched and with opposite arms of the hybrid junction ideally decoupled, a balanced amplifier represents a matched load to the signal generator. In the case of a mismatch, the stability conditions are given by the expression[96]

$$\gamma |\Gamma_1| |\Gamma_2| < 1, \quad (10.11)$$

where $|\Gamma_1|$ and $|\Gamma_2|$ are the moduli of the voltage coefficient of reflection from the signal source and from the amplifier load, respectively. Over a wide frequency band it is difficult to obtain good matching at all points of the band and the limitations imposed by (10.11) becomes significant. The limitations imposed on the balanced amplifier by its noise properties will be discussed later.

10.4. Traveling-wave amplifiers. Interest in traveling-wave amplifiers is due to their broadband characteristics. The most promising type of traveling-wave amplifier is one with a "very narrow strip diode" (see Fig. 38) located along the axis of a strip line and with two absorbing loads at a certain distance from the strip line[73]. The difficulty in realizing such a system lies in the fact that it amplifies in both directions. Slight mismatches in the line serve as sources of standing waves, and oscillation sets in if the gain is large. By virtue of the fact that the negative resistance acts at all frequencies, matching at the input and the output over a wide frequency band is hardly attainable. To eliminate oscillation in the distributed amplifier, a combined system was proposed, consisting of a dis-

tributed tunnel diode and a non-reciprocal ferrite attenuator [73,97].

Such a system amplifies signals passing in one direction and attenuates them in the other. The proposed device [97] (similar to that shown in Fig. 38) contains thin ferrite plates on both sides of a narrow strip diode, located along the axis of a strip line. The permanent magnet on which this device is mounted produces in the ferrite an inhomogeneous field, which is stronger near the axis and weaker towards the edges of the strip line. The frequency of the ferromagnetic resonance therefore decreases with increasing distance from the axis, maintaining the properties of the ferrite medium non-reciprocal over a very wide frequency band.

The difficulties connected with manufacturing distributed structures with negative resistance have not yet been overcome, and this has led to the development of a lumped-constant equivalent of a traveling-wave amplifier. In such a system the lumped tunnel diodes, discretely placed along the axis of a strip line, are separated by a ferrite medium, which absorbs the waves propagating in the undesired direction [98].

Another type of traveling-wave amplifier [99] (Fig. 45a) contains segments of transmission lines with lumped parameters, all having the same cutoff frequency. The junction capacitance and inductance of the tunnel diodes can be parts of such an artificial line. Each succeeding line segment (in the direction away from the generator) has a characteristic admittance larger than the preceding one. The negative admittances of the tunnel diodes connected at the terminals joining these lines provide the matching.

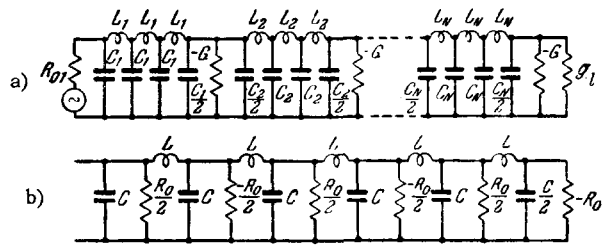

FIG. 45. Two types of traveling-wave amplifiers.

In the presence of N identical diodes with negative conductance −G, the conductance of the terminal load g_l should be

$$g_l = g_{01} + NG, \tag{10.12}$$

where g_{01} is the characteristic conductance of the line segment closest to the generator.

Such an amplifier has theoretically a large gain over a wide frequency band, the power gain being equal to the conductance ratio g_l/g_{01}.

In a somewhat different type of traveling wave amplifier [100] (Fig. 45b), all sections have the same characteristic resistance R_0. The last section of this amplifier is loaded by a negative resistance with absolute value $R = R_0$, and negative and positive resistances $-R/2$ and $R/2$ are connected in alternating sequence at each junction between sections. Each section is therefore loaded by a characteristic resistance amounting to either R or −R, the sign reversing in each section. The circuit has a theoretical bandwidth from zero to the limiting frequency of the tunnel diodes, and has a power gain equal to 2N, where N is the number of elements.

10.5. Stability of tunnel-diode amplifiers. Like all other regenerative amplifiers, the gain of tunnel-diode amplifiers exhibits a strong dependence on the load and generator impedances. This dependence is particularly large if the gain is large and can lead to self excitation in the limit. When the amplifier has a separate input and output, the sensitivity of the amplifier to a change in the input impedance is determined by the derivative of the gain with respect to the conductance, and amounts to [101]

$$\frac{\partial k_0}{\partial G_g''} = \frac{k_0}{G_g''}\left[1 - 2k_0^{1/2}\left(\frac{G_g''}{G_l''}\right)^{1/2}\right], \tag{10.13}$$

where k_0 is the nominal gain, determined by Eq. (10.2).

If the gain is made equal to $G_l''/4G_g''$, it will not be sensitive to changes in the external load. Unfortunately, the requirement $G_l'' \gg G_g''$ leads to an intolerable increase in the noise figure. The use of isolators and circulators greatly improves the stability of the amplifiers against changes in the external load. A change in the bias voltage can also lead to gain instability. The more linear the chosen working portion of the volt-ampere characteristic, the smaller the instability.

The temperature dependence of the negative conductance should also be taken into consideration in determining the stable gain of the amplifier.

The temperature coefficient of negative conductance is a nonlinear function of the bias voltage; definite portions on the negative branch have a weak temperature dependence [45,76]. However, a low noise figure is more likely to be the decisive factor in the choice of the bias voltage; consequently, if the amplifier is to be used in practice, gain stabilization may be necessary.

10.6. Noise in tunnel-diode amplifiers. The noise figure F is defined, as usual (see, for example, [102]), as the ratio of the total noise at the output terminals of the device (N_S) to that part of the noise, at the same terminals, which is due to the noise on a matched resistance external to the device at a temperature $T_0 = 290°K$ (N_{0S}):

$$F = \frac{N_s}{N_{0s}} = \frac{N_s}{kT_0\Delta f\gamma}. \tag{10.14}$$

The main source of noise in a tunnel diode is shot noise. Shot noise is connected with currents flowing through the p-n junction in both directions. Assuming that these currents are completely "uncorrelated,"

the equivalent saturation current I_{eq} (so named in analogy with the ordinary diode) is the sum of the aforementioned currents. When working on the descending portion of the volt-ampere characteristic, the current flowing through the junction in the opposite direction can be neglected as compared with the forward current. Consequently shot noise in a frequency interval Δf can be represented by a noise generator with mean square current [104]

$$\overline{i^2} = 2eI_{eq}\Delta f \approx 2eI\Delta f, \qquad (10.15)$$

where I is the measured diode current from the bias source.

To analyze the noise properties, real noise sources are frequently replaced by equivalent noise generators and ideal (i.e., non-noisy) resistances, as is done in Fig. 46, which shows the noise equivalent circuit of a tunnel diode [103-107]. In Fig. 46 the thermal noise of the resistance of the diode is replaced by an equivalent noise source e_s, with a mean-squared voltage

$$\overline{e_s^2} = 4kT_s\Delta f r_s, \qquad (10.16)$$

and the shot noise is represented by a current generator

$$\overline{i^2} = 4kT_d\Delta f g_e, \qquad (10.17)$$

where T_S and T_d are the absolute values of the temperatures of the loss resistance and of the p-n junction, and $g_e = eI/2kT_d$ is the equivalent noise conductance of the tunnel diode, defined by (10.15) and (10.17).

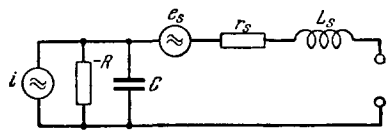

FIG. 46. Equivalent noise circuit of tunnel diode.

If $T_d = T_0 = 290°K$, then g_e (mho) = 0.02 I (mA).

The series and parallel equivalent noise circuits of an amplifier tuned to resonance, with separate input and output, are shown in Fig. 47.

In these circuits, the recalculated diode parameters are determined with the aid of expressions (5.3)—(5.5), and the parameters of the equivalent generators e′ and i″ are determined by equating the "exchangeable" powers [108]

$$\frac{\overline{i^2}}{-G} = \frac{4kT\Delta f g_e}{-G} = \frac{\overline{e'^2}}{-R'} = \frac{\overline{i''^2}}{-G''}. \qquad (10.18)$$

The concept of "exchangeable" power defined as the stationary value of the output power of the generator as a function of its output current or voltage, has the advantage over the customary value of "available" power in that it retains its meaning at both positive and negative values of the internal source resistance r_S (in the latter case the exchangeable power $P = |E|^2/4r_S$ is the power absorbed by the source from the external circuit).

The noise figure in an amplifier with separated input and output is

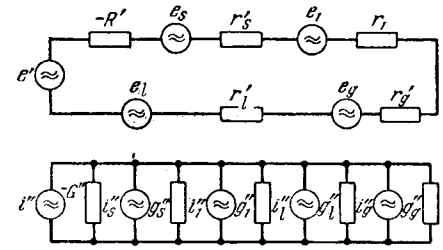

FIG. 47. Equivalent noise circuits of an amplifier tuned to resonance with separate input and output. a) Series circuit $\overline{e'^2} = 4kT - \Delta f r_e'$ where $r_e' = R'g_e/G$; b) tunnel circuit $i''^2 = 4kT\Delta f g_e''$, where $g_e'' = G''g_e/G$.

$$F = 1 + \frac{T_s}{T_0}\frac{r_s'}{r_g'} + \frac{T_1}{T_0}\frac{r_1'}{r_g'} + \frac{Tl}{T_0}\frac{r_l'}{r_g'} + \frac{T_d}{T_0}\frac{r_e'}{r_g'}$$

$$= 1 + \frac{T_s}{T_0}\frac{g_s''}{g_g''} + \frac{T_1}{T_0}\frac{g_1''}{g_g''} + \frac{Tl}{T_0}\frac{g_l''}{g_g''} + \frac{T_d}{T_0}\frac{g_e''}{g_g''}. \qquad (10.19)$$

If the gain is large and $T_S = T_1 = T_d = T_0$, this equation assumes the convenient form [91]

$$F = 1 + \frac{\frac{Ql}{Q_0} + \frac{r_l'}{r_g'}}{1 - \frac{Ql}{Q_0}} + \frac{1 + \frac{r_l'}{r_g'}}{\left(1 - \frac{Ql}{Q_0}\right)\left(1 - \frac{r_s}{R}\right)}$$

$$\times \frac{0.02\,IR + \frac{r_s}{R}(1 + \omega^2C^2R^2)}{\left(1 - \frac{\omega^2}{\omega_{lim}^2}\right)}, \qquad (10.20)$$

where I is in mA and R is in ohms.

The noise figure of a "reflection" amplifier with ideal circulator can be readily shown to be

$$F = \frac{1 + 0.02\,IR}{1 - \frac{r_s}{R}(1 + \omega^2C^2R^2)}. \qquad (10.21)$$

In the derivation of (10.21) it was assumed that the gain is large and all the losses are lumped in r_S. This equation determines the maximum noise figure attained at an operating frequency ω for specified diode parameters. We note that (10.21) is similar to the relations obtained in [103-106]. Thus, (10.21) can be represented as a function of three ratios of the tunnel diode parameters g_e/G, r_S/R, and ω/ω_{lim}: [106]

$$F = \frac{1 + \frac{g_e}{G}}{\left(1 - \frac{r_s}{R}\right)\left(1 - \frac{\omega^2}{\omega_{lim}^2}\right)}. \qquad (10.22)$$

If we neglect the losses in the resonant circuit and in the diode, then the noise figure away from the limiting frequency ω_{lim} is determined by the simple relation

$$F = 1 + 0.02IR, \qquad (10.23)$$

which is the limiting case of (10.21) and (10.22) (the same limit is reached for an amplifier with separated input and output if in addition $r_g' \gg r_l'$).

The noise figure of a balanced amplifier (Fig. 44) is determined by [109]

$$F = 1 + \gamma \frac{\Gamma_g^2}{1-\Gamma_g^2}\frac{T_l}{T_0} + \frac{4\frac{g_e}{G}G''g_0''(1+\Gamma_g^2\gamma)}{(1-\Gamma_g^2)(g_0''+G''-g_{\text{loss}}'')^2 T_0}\frac{T_d}{T_0}$$

$$+ \frac{4g_{\text{loss}}''g_0''(1+\Gamma_g^2\gamma)}{(1-\Gamma_g^2)(g_0''+G''-g_{\text{loss}}'')^2 T_0}\frac{T_1}{T_0}, \qquad (10.24)$$

where Γ_g^2 is the square of the modulus of the coefficient of voltage reflection from the input terminals of the balanced amplifier and γ is the reflection coefficient of each of the two identical amplifiers in the arms of the hybrid junction.

In this equation the second term is determined by the load noise entering the amplifier from the output side, while the third and fourth terms are due to the diode noise with account of the mismatch at the input. Under ideal conditions there is no reflection at the input and no losses in the tune circuit or in the diodes, so that the noise figure is

$$F = 1 + \left(1 - \frac{1}{\gamma_{12}}\right)\frac{eI}{2kT_0G}, \qquad (10.25)$$

where $\gamma_{12} = \gamma$ is the gain of the balanced amplifier in the absence of reflection at the input.

This equation shows that the noise figure of an ideal balanced amplifier increases with increasing gain and reaches in the limit the value 0.02 IR, which characterizes a single amplifier of the "reflecting" type. The presence of a mismatch at the input, as can be seen from (10.24), increases the noise figure of the balanced amplifier. Finally, the overall noise figure of the system will be appreciably increased, in proportion to the mismatch at the input and to the gain of the amplifier, by noise from the succeeding stages. The rise in the effective noise temperature T_{bal} of the balanced amplifier above to the minimum effective noise temperature T_{min} of the component amplifiers, resulting from all these factors, can be expressed by the approximate relation[110]

$$\frac{T_{bal}}{T_{min}} \approx 1 + \Gamma_g^2\gamma\left(1 + \frac{T_{e2}}{T_{e\,min}}\right). \qquad (10.26)$$

Here T_{e2} —noise temperature at the input to the second stage.

It is easy to see that mismatching greatly deteriorates the performance of a balanced amplifier. The use of circulators negates the entire idea of this circuit, for then one can make a single reflecting amplifier do. The applications of balanced amplifiers are therefore quite limited.

The noise figure can be reduced by decreasing the gain and increasing g_0''. For a reflecting amplifier with a circulator* we have

$$F = 1 + \frac{4T}{T_0}g_0''G''\frac{\frac{eI}{2kTG} + \frac{g_{\text{loss}}''}{G''}}{(g_0'' - g_{\text{loss}}'' + G'')^2}, \qquad (10.27)$$

*The authors of [96,107,109] do not investigate the frequency variation of the noise figure and consequently the expressions analogous to (10.24), (10.27), and (10.28) do not contain the transformed values of the conductances.

where $T = T_1 = T_S = T_d$. The need for eliminating the noise in the succeeding stages should, however, lead to the use of several series-connected amplifiers. It can be shown[107] that the overall noise figure of such an n-stage amplifier of the "reflecting" type with large overall gain ($\gamma^n \to \infty$), in which each stage has a low gain γ and is "decoupled" from the others, is

$$F_n = 1 + \frac{T}{T_0}\left[\frac{1-\frac{1}{\gamma^n}}{1-\frac{g_{\text{loss}}''}{G''}}\right]\left(\frac{eI}{2kTG} + \frac{g_{\text{loss}}''}{G''}\right) \qquad (10.28)$$

and can be made equal to the noise figure of a high-gain stage. (This entails only an increase in the bandwidth of the system.)

These equations show that in a correctly designed amplifier the main contribution to the noise factor is made by the shot noise. To reduce the noise figure it is essential to decrease the losses in the diode and in the resonant circuit of the amplifier and to operate away from the limiting frequency. If the latter condition is not satisfied the noise from the series resistor will adversely affect the noise properties of the amplifier. Strong coupling with the generator and weak coupling with the load are furthermore necessary conditions for a reduction in the noise figure of an amplifier with separated input and output. The limiting value of the noise figure for all three types of amplifiers considered above is given by (10.23), so that the product IR can be taken as a measure of the noise performance of tunnel diodes. A typical IR vs. bias voltage curve is shown in Fig. 48 together with the volt-ampere characteristic.

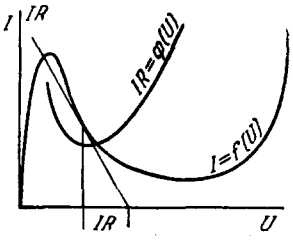

FIG. 48. Typical dependence of the product IR of a tunnel diode on the bias voltage.

It is seen from Fig. 48 that excessive current leads to an increase in the modulus of the negative resistance; in this connection, the minimum value of IR is usually in the central portion of the descending part of the volt-ampere characteristic.

No theoretical limits have yet been established for $(IR)_{min}$. We can note only that the possibility of decreasing $(IR)_{min}$ is connected with the use of diodes having narrow forbidden bands and low effective masses[104]. Thus, $(IR)_{min}$ is 0.12 V for gallium arsenide and 0.06 V for germanium. Worthy of attention is the use of gallium antimonide, for which $(IR)_{min}$ = 0.043 V;*[111] the noise figures corresponding to the three indicated values of IR are 5.3, 3.4, and 2.7 dB.

*Oscillations with fundamental frequency up to 62.5 Gc were produced with GaSb diodes.[111]

10.7. Experimental results. The technology used to produce tunnel-diode amplifiers for the microwave band is similar in many respects to the above-described methods for the design of oscillators. This pertains to the construction of cavities, connecting lines, and bias circuits. At the lower microwave frequencies coaxial and strip lines are used for the most part, while at higher frequencies waveguides and cavities are employed.

The first description of the tunnel-diode amplifier was published in 1959 by Chang [112] who constructed an amplifier for 30—80 Mc with a noise figure 4.5 dB.

Another amplifier with separate input and output operating at 405—465 Mc, with a gain 15 dB and a bandwidth $\Delta f = 12$ Mc, had a noise figure of 4.2 dB. The amplifier consisted of a short-circuited strip cavity, one end connected to a ZJ56 diode (made by General Electric; G = 0.01 mho, C = 7 pF, $L_S = 6 \times 10^{-9}$ H, $R_S \approx 1\ \Omega$) strongly coupled to the generator and weakly coupled to the load. The gain changed by 3 dB as the SWR changed from 1 to 1.2 without an oscillator and from 1 to 1.8 with an oscillator. [101]

An amplifier of the "reflecting" type with circulator operated at 4500 Mcs with $\sqrt{\gamma}\,\Delta f = 275$ Mc, $\gamma = 23$ dB, and $\Delta f = 20$ Mcs. The noise figure of the amplifier was 7 dB ($\sqrt{\gamma}\,\Delta f = 360$ Mc, F = 8.6 dB in a different version) [113,114]. The coupling with the cavity containing the diode was by means of a short section of waveguide operating beyond cutoff. The product $\sqrt{\gamma}\,\Delta f$ was constant as the gain varied from 17 to 25 dB. Comparison of the experimental and calculated values of $\sqrt{\gamma}\,\Delta f$ helped determine the cavity and diode losses and calculate the amplifier noise figure. The noise figure so calculated was close to the experimental values.

A stable reflecting type amplifier with circulator was designed for a gain of 36—38 dB at 6.8, 10.8, and 25.8 Gc [115]. The amplifier resonant circuit was a conical cavity shunted by a welded tunnel diode made of GaAs. The noise figure was 7.5, 10.7, and 11.5 dB at 6.8, 10.7, and 11.5 Gc, respectively.

Amplification at even higher frequencies from 55 to 85 Gc, was obtained [111] with welded diodes of n-type GaAs [85] built into a tapered 3.1 × 0.25 mm waveguide, in analogy with [88]. The amplifier was of the reflecting type with circulator and its parameters at 55 Gc were $\gamma = 35$ dB, $\sqrt{\gamma}\,\Delta f = 400$ ($\gamma = 20$ dB), and F = 16—18 dB. The noise figure exceeded the values given by (10.23) (6—7 dB for the diode employed) because of the large losses in the resonant circuit of the amplifier.

The RCA type SS500 amplifier has a minimum gain of 15 dB in the 1275—1325 Mc band and a noise figure of 6 dB. The input saturation power of the amplifier is 30 microwatts, and the operating drain is 10 mA at 0.1 V. [89]

The minimum attained noise figure at 400 Mc was 2.7 dB ($T_{eff} = 250°K$) [116].

Thus, the experimental results confirm the theoretical feasibility of amplification with the aid of tunnel diodes in the centimeter and millimeter bands. Reduction of the noise figure to values governed by the shot noise is apparently the problem of the near future. Further lowering of the shot noise is connected with the problem of reducing the excess current.

But even in the present form the use of tunnel-diode amplifiers is increasing, since there are many applications where their exceeding simplicity is more than sufficient compensation for the somewhat higher noise as compared with parametric amplifiers.

11. OTHER APPLICATIONS OF TUNNEL DIODES

Microwave applications of tunnel diodes are not confined to amplifiers and oscillators. They can be used for frequency conversion (mixing), detection, and application in the superregenerative mode. These applications are based on the use of the negative resistance in the nonlinearity of the volt-ampere characteristic of the tunnel diode.

11.1. Tunnel-diode converter (mixer). The equivalent circuit shown in Fig. 49 consists of three tuned circuits coupled by the tunnel diode [117]. The circuit can operate with an external heterodyne (at frequency ω_3) as a mixer, converting the signal frequency ω_1 into an intermediate frequency ω_2. If the self-excitation conditions are satisfied at the frequency ω_3, the circuit operates as a converter without an external heterodyne.

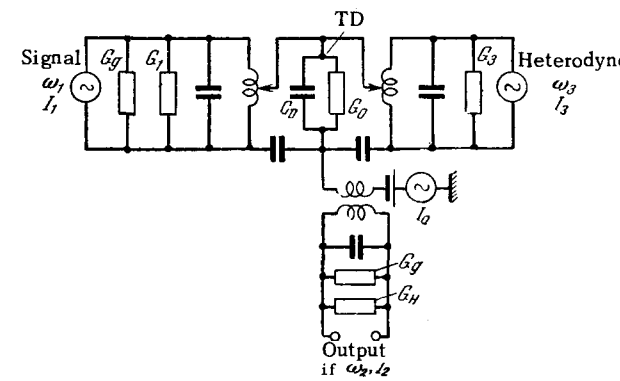

FIG. 49. Equivalent circuit of tunnel-diode converter. TD – tunnel diode.

The operating point is chosen on the rising branch of the volt-ampere characteristic near the current maximum. The working portion of the characteristic is approximated by the quadratic equation

$$I = g_0 U - AU^2, \qquad (11.1)$$

where g_0 is the conductance at the operating point and A is a proportionality coefficient.

By analyzing the equations for the loop currents at the three frequencies (ω_1, ω_2, ω_3), Chang [117] has shown that a conversion coefficient greater than unity

is obtained if the amplitude of the voltage U_3 at the heterodyne frequency is sufficient to satisfy the condition $AU_3/g_0 > 1$. By conversion coefficient is meant the ratio of the output power of the intermediate frequency to the maximum power delivered by the signal source.

Relations for the conversion coefficient, bandwidth, and noise figure of the mixer were derived [117] by assuming that (11.1) holds. The experimental mixer described in that paper converted 210 Mc signals into 30 Mc if signals. The heterodyne frequency was 240 Mc. Using a large-current germanium tunnel diode (I_p = 35 mA, g_0 = 0.21 mho), a conversion coefficient of 6.0 dB was obtained with a bandwidth 0.9 Mc and a noise figure 5.2 dB. The corresponding values for a gallium-arsenide diode were 22.7 dB, 0.15 Mc and 2.8 dB. These experimental data agreed with the calculation [117].

Promising results were obtained at a signal frequency of 1200 Mcs (f_{het} = 1170 Mcs, f_{if} = 30 Mcs).[118] Unlike the mixer described above, the operating point was chosen on the descending branch of the characteristic, and the peak current was much smaller (I_p = 1 mA). It was therefore not at all difficult to satisfy the dc stability conditions. The authors obtained a high conversion coefficient in a 4 Mc band and a 4.5 dB noise figure which is close to theoretical (with exact tuning the results were even better but the conversion was unstable).

Breitzer [119] analyzed the operation of the mixer using an analytic expression similar to (3.3) for the volt-ampere characteristic, with account of the current flowing in both directions through the p-n junction. Assuming these currents to be amplitude-modulated by the heterodyne and assuming inverse conversion, he arrived at conclusions that differed somewhat from those of Chang. His report does not contain experimental data, but the results of Green and Sard [120] are closer to Breitzer's calculations than to Chang's conclusions. The available data indicate that tunnel diodes offer certain promise for use in frequency conversion.

11.2. Tunnel diode detector. A detector of very high sensitivity can be realized with the aid of a tunnel diode in a circuit that combines regeneration and detection. A sensitive detector of this type was briefy described by Montgomery [121]. The equivalent circuit of Fig. 26 can be used. The maximum sensitivity was observed with the operating point located near the oscillation point. The frequency dependence of the sensitivity for several values of the negative resistance $-R$ is shown in Fig. 50, where R is normalized relative to R_0, the negative resistance corresponding to a start of oscillation. The best short-circuit current sensitivity of the unit amounted to 10^5 A/W and the minimum detected signal level was 140 dB below 1 watt. (By current sensitivity is meant the ratio of the rectified short-circuit current to the maximum power of the signal

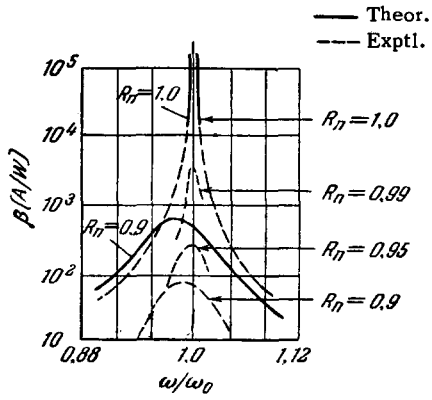

FIG. 50. Dependence of the short-circuit current sensitivity on the frequency in a regenerative detector.

source). The bandwidth near the oscillation point, however, was small and the stability insufficient. A high stable sensitivity, 5×10^2 A/W, corresponding to a detected power of 120 dB/W, occurred at a bandwidth up to 2%. The amplitude characteristic was quadratic up to 60 dB-mW. The dynamic range was limited on the high-power side by the occurrence of oscillation.

The authors of [123] recommend the use of a so-called "backward" tunnel diode as a detector. Such a tunnel diode is characterized by the fact that a degenerate semiconductor is located only on one side of the p-n junction. Consequently the forward branch of its volt-ampere characteristic is similar to that of an ordinary diode. The current flowing through the p-n junction in the backward direction is due to the tunnel effect. The volt-ampere characteristic of the backward diode has a weak temperature dependence and a sharply pronounced inflection near the origin. At the point of inflection such a diode can have a large ratio I''/I' of the second derivative of the current with respect to the voltage to the first derivative.

Thus, for example, a value $I''/I' = 1.23$ e/kT was obtained for gallium arsenide diodes at a bias of 5 mV, and 2.17 e/kT at a bias of 1 mV [123] ("ordinary" crystal detectors cannot have I''/I' larger than 0.5 e/kT.)

11.3. Superregenerator using a tunnel diode. As is well known, the superregenerative amplifier offers very high sensitivity with sufficient stability, whereas the stability of the regenerative decreases at the limiting values of the gain.

The operating point of a tunnel diode used as a superregenerator [124] is biased into the negative-resistance region, and the oscillation conditions are satisfied in the resonant circuit. The amplitude of the auxiliary source should be sufficient for periodic interruption of these oscillations. It is possible to combine in a single circuit a superregenerative amplifier and the auxiliary-oscillation generator, and also a superregenerative amplifier and detector. The equivalent circuit of a superregenerative detector for 6.25 Mc is shown in Fig. 51 (auxiliary-source fre-

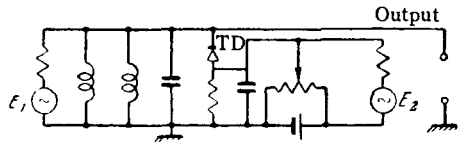

FIG. 51. Equivalent circuit of superregenerative detector. TD – tunnel diode.

quency 0.1 Mc).[124] Report of regeneration at a frequency 980 Mc confirms the feasibility of circuits of this type in the microwave band.[121]

12. USE OF TUNNEL DIODES FOR PHYSICAL RESEARCH

The quantum-mechanical nature of the tunnel diode affords experimenters an opportunity to investigate directly the interaction between electrons and phonons.

The first such interactions were observed in [48], where characteristic kinks in the forward branch of the volt-ampere characteristic of tunnel diodes were observed at helium temperatures (Fig. 52). The voltages at which these kinks occur corresponded exactly to the energies of the phonons existing in the semiconductor. In germanium doped with antimony, all four acoustic phonons were observed, two longitudinal and two transverse.

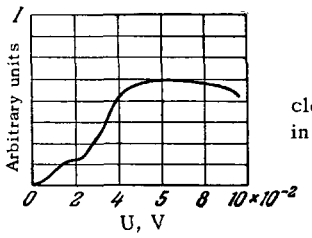

FIG. 52. Experimental curves disclosing the participation of the phonons in the tunnel effect.

In germanium doped with phosphorus and arsenic, no kinks were observed on the characteristic. This indicates either direct tunneling in such crystals, or indirect tunneling with scattering on impurities. The latter is the more probable. This result is confirmed also by the form of the temperature dependence of the tunnel current (see Sec. 5.2) and measurements of the backward branch of the volt-ampere characteristic.[127]

Electron-phonon interactions were observed also in silicon tunnel diodes[48]. Investigations of this type are beginning to assume a quantitative character. Hall has shown[59] that a volt-ampere characteristic with kinks can be resolved into a sum of several similar curves (see dashed curves in Fig. 53a). The curves are shifted along the voltage axis by amounts proportional to $\hbar\omega_i$ corresponding to the phonons. Their amplitudes differ because of the different degree of phonon-electron interaction. A clearer picture is obtained by plotting the second derivative of the tunneling current with respect to the voltage (Fig. 53b)[59].

A careful study has shown that each of the dashed curves of Fig. 53a falls off towards the voltage axis

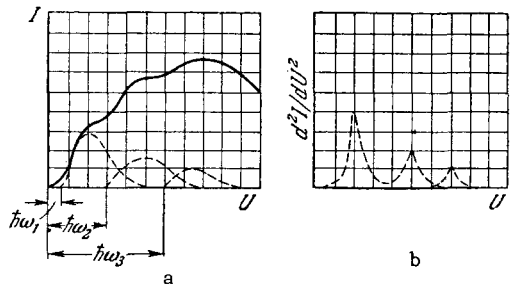

FIG. 53. Curves illustrating the participation of phonons in the tunnel effect.

not steeply, but asymptotically. This phenomenon is interpreted in [59] as a consequence of the smearing of the lower boundary of the conduction band inside the forbidden band of the semiconductor.

Low-temperature investigations of the tunnel current of diodes made of semiconductor compounds (of the type $A^{III}B^{V}$, etc.) have made it possible to observe interactions between electrons and polarons.[51]

Investigations of the volt-ampere characteristic of the tunnel diode in strong magnetic field[125] have yielded the effective mass of the charged carriers. Such a determination is based on the well known de Haas–van Alphen effect. In tunnel diodes this effect manifests itself in oscillations of the current flowing through the diode. Chynoweth[59], in a study of tunnel diodes made of PbTe in fields exceeding 40,000 G, registered nine such oscillations and more. By measuring their period, he was able to find the value of the Bohr magneton, namely

$$\mu = \frac{eh}{m^* c} \qquad (12.1)$$

and consequently the effective mass m*. An investigation of the behavior of tunnel diodes in a magnetic field is apparently also of practical interest, in view of the decrease in I_p/I_V with increasing field intensity[125,128].

The measurement of forward and backward currents in the tunnel diode enables us to estimate the probability of the tunneling, and consequently, to estimate all the quantities contained in the expressions defining this probability, (2.3) and (2.4).

To this end, the width of the forbidden band E_g was measured in [52,126] as a function of the hydrostatic pressure acting on the diode.

Finally, in Sec. 5.3 we already mentioned tunnel spectroscopy, which makes possible a study of deep traps in the forbidden band of the crystal.

The possibilities of tunnel diode applications are far from exhausted by the cited examples. The vigorous study of these diodes, now in progress, will undoubtedly yield in the nearest future valuable information on the physical phenomena in solids.

[1] J. I. Frenkel and A. F. Ioffe, Physik. Z. Sowjetunion 1, 60 (1932).

[2] W. Shockley, Electrons and Holes in Semiconductors, N. Y., Van Nostrand, 1950.

[3] B. M. Vul, FTT 2, 2961 (1960), Soviet Phys. Solid State 2, 2631 (1962).
[4] A. G. Chynoweth, Progress in Semiconductors 4, 95 (1960).
[5] L. Esaki, Phys. Rev. 109, 603 (1958).
[6] É. V. Shpol'skiĭ, Atomnaya fizika (Atomic Physics), vol. 1, Gostekhizdat, 1950.
[7] I. A. Lesk, N. Holonyak, et al., IRE Wescon Conv. Record. Pt. 3, 9 (1959).
[8] C. Zener, Proc. Roy. Soc. 145, 523 (1934).
[9] K. B. McAffee et al., Phys. Rev. 83, 650 (1951).
[10] Hartman, Michelitsch, and Steinhäuser, Arch. d. elektr. Übertragung 15, 3, 125 (1961).
[11] W. Franz and L. Tewordt, Halbleiterprobleme, Bd. III, Braunschweig, 1956.
[12] L. V. Keldysh, JETP 33, 994 (1957), Soviet Phys. JETP 6, 763 (1958).
[13] L. V. Keldysh, JETP 34, 962 (1958), Soviet Phys. JETP 7, 665 (1958).
[14] Hall, Bardeen, and Blatt, Phys. Rev. 95, 559 (1959).
[15] C. W. Bates, Phys. Rev. 121, 1070 (1961).
[16] I. I. Ivanchik, FTT 3, 103 (1961), Soviet Phys. Solid State 3, 75 (1961).
[17] P. J. Price and J. H. Radcliff, IBM J. 3, 364 (1959).
[18] E. O. Kane, J. Appl. Phys. 32, 83 (1961).
[19] V. L. Bonch-Bruevich, Radiotekhnika i élektronika (Radio Engineering and Electron Physics) 5, 2033 (1960).
[20] F. A. Trumbore and A. A. Tartaglia, J. Appl. Phys. 29, 1511 (1958).
[21] L. Esaki, Solid State Physics in Electronics and Telecommunications, Vol. I, Acad. Press., London-New York, 1960.
[22] E. Spenke, Electronics Semiconductors, McGraw-Hill, N.Y., 1958.
[23] K. S. Shifrin, ZhTF 14, 40 (1944).
[24] P. P. Debye and E. M. Conwell, Phys. Rev. 93, 693 (1954).
[25] F. A. Trumbore, Bell System Tech. J. 39, 205 (1960).
[26] N. Holonyak and I. A. Lesk, Proc. IRE 48, 1405 (1960).
[27] R. Gremmelmaier, Z. Naturforsch. 14a, 1072 (1959).
[28] R. L. Batdorf, J. Appl. Phys. 31, 613 (1960).
[29] F. A. Trumbore and E. M. Porbansky, J. Appl. Phys. 31, 2068 (1960).
[30] R. A. Smith, Semiconductors, Cambridge, 1959.
[31] H. S. Sommers, Proc. IRE 47, 1201 (1959).
[32] R. N. Hall, Proc. IRE 40, 1512 (1952).
[33] R. F. Rutz, IBM J. 3, 372 (1959).
[34] V. I. Fistul' and I. D. Abezgauz, Trans. Conf. on Impact Ionization and Tunnel Effect in Semiconductors, AN AzerbSSR, Baku, 1961 (in press).
[35] R. N. Rubinshteĭn and V. I. Fistul', Zav. laboratoriya (Plant Laboratory) 27, 1242 (1961).
[36] M. Michelitsch, Naturwiss. 47, 274 (1960).
[37] Collection of Translations, "Germanium," ed. by D. A. Petrov, IL, 1955, pp. 218-230.
[38] H. S. Sommers et al., IRE Wescon Conv. Record. Pt. 3, 3 (1959).
[39] R. C. Sims, E. R. Beck et al., Proc. IRE 49, 136 (1961).
[40] E. G. Nielsen, Proc. IRE 48, 1903 (1960).
[41] H. S. Sommers, Solid State Conference, Cornell University, 1959.
[42] N. A. Belova and A. N. Kovalev, see [19] 6, 160 (1961).
[43] Y. Furukawa, J. Phys. Soc. Japan 15, 730 (1960).
[44] B. G. Zhurkin, see [34], in press.
[45] W. Cady, Trans. Colloque International sur les Dispositifs a Semiconducteurs, Paris, 20—25 Fevrier, 1961.
[46] R. N. Hall, IRE Transactions ED7, 1 (1960).
[47] Vul, Shotov, and Grishechkina, FTT 3, 667 (1961), Soviet Phys. Solid State 3, 489 (1961).
[48] N. Holonyak, I. A. Lesk et al., Phys. Rev. Lett. 3 (4), 167 (1959).
[49] B. Lax, Revs. Modern Physics 30, 122 (1958).
[50] T. Yajima and L. Esaki, J. Phys. Soc. Japan 13, 1281 (1958).
[51] R. Hall, Trans. Int. Conference on Semiconductors, Prague, 1960.
[52] S. L. Miller et al., Phys. Rev. Lett. 4, 60 (1960).
[53] R. A. Logan and A. G. Chynoweth, Bull. Amer. Phys. Soc. 5, 160 (1960).
[54] A. G. Chynoweth et al., Phys. Rev. 121, 684 (1961).
[55] R. E. Blair and J. W. Easley, J. Appl. Phys. 31, 1772 (1960).
[56] E. O. Kane, Trans. Int. Conference on Semiconductors, Prague, 1960.
[57] W. Franz, Handb. d. Phys. Bd. 17, Berlin, 1956, p. 155.
[58] N. Holonyak, J. Appl. Phys. 32, 130 (1961).
[59] Materials of Conference on Tunnel Diodes, USA, 1959.
[60] Andronov, Vitt, and Khaĭkin, Teoriya kolebaniĭ (Theory of Oscillations), Fizmatgiz, 1959.
[61] U. S. Davidson et al., Electronic Design, February 3, p. 17 (1960).
[62] A. G. Chynoweth, Phys. Rev. 120(5), 1620 (1960).
[63] Fukui, Denki, Zusin, Gakkai, and Dyassi 43, 1351 (1960).
[64] Electronic news 199, 22 (1960).
[65] V. I. Fistul', op. cit. [19], in press.
[66] Dermit, Lockwood, and Hauer, Proc. IRE 49, 519 (1961).
[67] NTZ, Bd. 13, Hf. 4, 191, 3—7 (1960).
[68] Electronic news 4, No. 175, 13 (1959); 4, No. 176, 41 (1959).
[69] L. Esaki, Paper at XII General Assembly, London, 1960.
[70] H. Bode, Network Analysis and Feedback Amplifier Design, Van Nostrand, N.Y., 1945.

[71] M. A. Aĭzerman, Lektsii po teorii avtomaticheskogo regulirovaniya (Lectures on the Theory of Automatic Control), Fizmatgiz, 1958.
[72] Wallace, see [113].
[73] M. E. Hines, Bell System Tech. J. **39**, 477 (1960).
[74] H. G. Dill et al., Electronics **33**, 5, 62 (1960).
[75] A. M. Gudman, Rev. Sci. Instr. **31**, 286 (1960).
[76] T. O. Krueger, Proc. Nat. Convent on Military Electr., 4th, 1960, p. 203.
[77] U. S. Davidson et al, Proc. IRE **49**, 372 (1961).
[78] H. G. Dill, Electronic Design, January 18, p. 30 (1961).
[79] D. E. Nelson and F. Sterzer, IRE Wescon Conv. Record, Pt. 1 (1960).
[80] C. Brunneli, Proc. IRE **25**, 1595 (1937).
[81] F. Sterzer and D. E. Nelson, Proc. IRE **49**, 744 (1961).
[82] J. K. Pulfer, Proc. IRE **48**, 1155 (1960).
[83] E. L. R. Webb and J. K. Pulfer, Canadian Electr. Eng. **55**, 40 (1961).
[84] G. Dermit, Proc. IRE 49, 1033 (1961).
[85] C. A. Barrus, J. Appl. Phys. **32**, 1031 (1961).
[86] F. Sterzer, Proc. IRE **47**, 1317 (1959).
[87] R. Trambarulo and C. A. Barrus, Proc. IRE **48**, 1776 (1960).
[88] C. A. Barrus, Proc. IRE **48**, 2024 (1960).
[89] Electronic Design **9**, May 10 (1961).
[90] Microwave J. **4**(2), 118 (1961).
[91] L. A. Birger, op. cit. [19] **4**(11), 1894 (1961).
[92] E. S. Kuh and J. D. Patterson, Proc. IRE **49**, 1043 (1961).
[93] J. S. Logan Proc. IRE **49**, 832 (1961).
[94] E. W. Sard, Proc. IRE **49**, 357 (1960).
[95] S. H. Autler, Proc. IRE **46**, 1880 (1958).
[96] J. J. Sie, Proc. IRE **48**, 1321 (1960).
[97] W. W. Anderson and M. E. Hines, Trans. IRE, MTT9, No. 1, 63 (1961).
[98] Electronics **33**, March, 25, p. 84 (1960).
[99] N. F. Moody and A. G. Wacker, Proc. IRE **49**, 835 (1961).
[100] C. A. Skalski, Proc. IRE **48**, 1909 (1960).
[101] G. Schaffner, IRE Wescon Conv. Record, Pt. 2, 86 (1960).
[102] J. J. Sie and S. Weisbaum, IRE Nat. Conference, March 23—26, 1959, p. 141.
[103] A. Van der Ziel, Proc. IRE **48**, 1321 (1960).
[104] J. J. Tieman, Proc. IRE **48**, 1418 (1960).
[105] P. Penfield, Proc. IRE **48**, 1478 (1960).
[106] E. G. Nielsen, Proc. IRE **48**, 1903 (1960).
[107] M. E. Hines and W. Anderson, Proc. IRE **48**, 789 (1960).
[108] H. A. Haus and R. B. Adler, Proc. IRE **46**, 1517 (1958).
[109] M. Feist, Proc. IRE **49**, 975 (1961).
[110] J. C. Grene, Proc. IRE **49**, 626 (1961).
[111] C. A. Burrus, Proc. IRE **49**, 1961 (1961).
[112] K. K. N. Chang, Proc. IRE **47**, 1268 (1959).
[113] A. Yariv and J. S. Cook, Proc. IRE **49**, 739 (1961).
[114] A.Yariv et al., Proc. IRE **48**, 1155 (1960).
[115] R. F. Trambarulo, Proc. IRE **48**, 2022 (1960).
[116] G. Wade, Proc. IRE **49**, 880 (1961).
[117] K. K. N. Chang et al., Proc. IRE **48**, 854 (1960).
[118] L. E. Dickens and C. R. Gneiting, IRE Trans. MMT9, 99 (1961).
[119] D. J. Breitzer, Proc. IRE **48**, 935 (1960).
[120] J. C. Grene and E. W. Sard, Proc. IRE **49**, 350 (1961).
[121] M. D. Montgomery, Proc. IRE **49**, 826 (1961).
[122] Semiconductor products, September, 1960.
[123] J. Karlovsky and A. Marek, Czech. Phys. J. **B11**, 76 (1961).
[124] A. G. Jordon and R. Elco, Proc. IRE **48**, 1902 (1960).
[125] A. G. Chynoweth et al., Phys. Rev. Lett. **5**, 548 (1960).
[126] M. J. Nathan and S. L. Miller, Bull. Amer. Phys. Soc., Ser. II, **5**, 265 (1960).
[127] J. V. Morgan and E. O. Kane, Phys. Rev. Lett. **3**, 466 (1959).
[128] A. R. Calawa et al., Phys. Rev. Lett. **5**, 55 (1960).
[129] Coll. "Tunnel Diodes" (Translations), ed. V. I. Fistul', IL, 1961.
[130] Microwave J. **4**(5), 114 (1960).

Translated by J. G. Adashko