PUBLIKACIJE ELEKTROTEHNIČKOG FAKULTETA UNIVERZITETA U BEOGRADU
PUBLICATIONS DE LA FACULTÉ D'ÉLECTROTECHNIQUE DE L'UNIVERSITÉ À BELGRADE

SERIJA: MATEMATIKA I FIZIKA — SÉRIE: MATHÉMATIQUES ET PHYSIQUE

№ 245 (1969)

THE STUDY OF THE ELECTRON TRANSPORT PHENOMENA IN SEMICONDUCTOR P-N STRUCTURES*

Dimitrije A. Tjapkin

PREDGOVOR

Ovaj rad predstavlja skraćenu i neznatno modifikovanu verziju doktorske disertacije autora (ref. [21]).

Predmet ovog rada su istraživanja nekih zakonitosti kod elektronskih transportnih procesa u nehomogenim poluprovodnicima, odnosno P-N prelazima. *Cilj* je da se ustanove kvantitativne ili polukvantitativne zavisnosti za one procese koji do sada nisu bili dovoljno detaljno razmotreni.

Razmotreni su neki osnovni problemi teorije P-N prelaza (Gl. 2) i *granični režimi* pri direktnoj i inverznoj polarizaciji (Gl. 3), s tim što su svi bitni rezultati i verifikovani na odgovarajući način (Gl. 4). Naročita pažnja posvećena je *temperaturskim zavisnostima.*

Pored toga, težilo se da svi problemi budu razmotreni sa svoje *osnovne*, a manje sa aplikativne strane. *Koncepcija* koje smo se držali, bila je ta da se manje išlo za matematički tačnim rešenjima ili numeričkim postupcima a više za prostijim ali fizički adekvatnijim rešenjima koja omogućuju suštinsku analizu pojava.

Lista oznaka data je u prilogu A. Svuda je korišćen internacionalni sistem (I.S. — odnosno MKSA sistem, relacije se zapisuju u racionalizovanom obliku).

Jedan deo ovog rada predstavlja rezultate koji su postignuti zahvaljujući pomoći Saveznog fonda za naučni rad (Ugovor br. 3-1692/1 sa Ins. za fiziku), tako da posebnu zahvalnost dugujem Institutu za fiziku i Elektrotehničkom fakultetu u Beogradu za materijalnu pomoć pri ovom radu, odnosno za stavljanje na raspolaganje svojih istraživačkih kapaciteta. Takođe se zahvaljujem svim mojim kolegama sa kojima sam, tokom rada na tezi, vodio vrlo korisne diskusije o određenim problemima.

PREFACE

This work represents a shortened and slightly modified version of the author's Dissertation submitted to the Faculty of Electrical Engineering of the Belgrade University (see Ref. [21]).

The *subject matter* of this work is the investigation of some phenomena encountered in the electron transport processes in inhomogeneous semiconductors i.e. P-N transitions. The main aim is to establish in more or less quantitative terms the relationships for those processes which have not yet been studied to any appreciable extent.

Some basic problems concerning the P-N *transition theory* (Chapt. 2) and the *limiting regimes* for the direct and reverse polarizations (Chapt. 3) have been discussed and the most

* Received February 6, 1969

important results verified adequately (Chapt. 4). Special attention is given to *temperature dependence* of some basic parameters.

Throughout the work emphasis is put more on the fundamental aspects of the processes involved, rather than on the possible applications. In dealing with these problems, we have also been guided by the idea that simple but physically more adequate approach may lead to better understanding of the essential concept than the analysis based on more exact numerical computational methods.

List of symbols is given in the Appendix A. The international system (MKSA rationalized system) is used throughout.

The work reported herein was partially sponsored by the Yugoslav Federal Fund for Scientific Research (under Contract No. 3-1692/1 with Institute of Physics — Belgrade). My special thanks are due to the Institute of Physics and to the Faculty of Electrical Engineering in Belgrade for their financial support, i.e. for giving me the possibility to use their research capacities. I also feel grateful to all my colleagues for their very useful discussions.

CONTENTS

Chapter 1. — INTRODUCTION

This problem derives its origin from some papers published at the end of the last century and afterwards at the beginning of this century ([1], [2], [3], and others — the earliest period: natural crystals PbS, Fe, CuO, Cu_2O — rectifiers). It attained its full development in the thirties ([4], [5], [6], and others — the period of the theory formation, the first systematic studies and the hypothesis of rectifier action; Cu_2O, PbS, SiC, Se — rectifiers and detectors) and subsequently in the forties (the final theory formation done by Davydov [7], Mott in 1938, Schottky and Spenke in 1939, Pekar in the year 1940, Bethe in 1942, Shockley in 1949 and others; the notions of holes and injection; modern singlecrystal semiconductors Ge, Si, ...), after which there was a certain stagnation. Apart from scientific interest (the problem is examined here from that point of view) this problem has become again important because of the appearance of numerous new applications as in microelectronic integrated circuits, field-effect transistors, cryo-electronics, semiconductor lasers and others.

Intermetalic compounds (III—V) — single-crystals InSb, GaAs, GaP and others (for instance SiC), as well as oxide films (for instance SiO_2) have received much attention in present days (see also Table T1.1). More quantitative results have been obtained by various authors and new physical interpretations of some phenomena have been presented (for instance [16] — dealing with physical interpretation of saturation current phenomenon in field effect devices). Rectifying and transistor effects have now also been treated extensively in the current textbooks and monographies ([8]—[15] and many others), among which especially [8], [9] and [10] should be pointed out.

Fig. 1.1. illustrates the *definition of a heterogeneous and inhomogeneous solid* (junction), from which it can be seen clearly that the latter implies inhomogeneity only in respect to impurity distribution (N and P being the concentrations of donors and acceptors, respectively; n_i — intrinsic concentration of carriers), while all others „intrinsic" properties (a — lattice constant, m^* — effective mass, ε — dielectric constant and so on) are the same for $x>0$ and for $x<0$.

Some specific problems concerning the electronic transport processes in the space charge region (Chapter 2) will be examined with the particular emphasis put upon the regime of stronger fields and currents. The temperature dependence of various relevant parameters will also be discussed.

However our discussion will be devoted for the most part to the case of *depleted* space charge region, for which, from well-known relation for the space charge density

$$\rho = en_i \mathcal{N} + (p-n)\, e + en_i \mathcal{N}_R \tag{1.1a}$$

we have

$$|n_i \subset\!\!\!\supset \mathcal{N}| > \text{or} \gg |(p-n)|. \tag{1.1 b}$$

Apart from that, all donor and acceptor centres are considered to be ionized*), which is not true for recombination centres (concentration $N_R \equiv n_i \subset\!\!\!\supset \mathcal{N}_R$). One-dimensional (planar) case is treated.

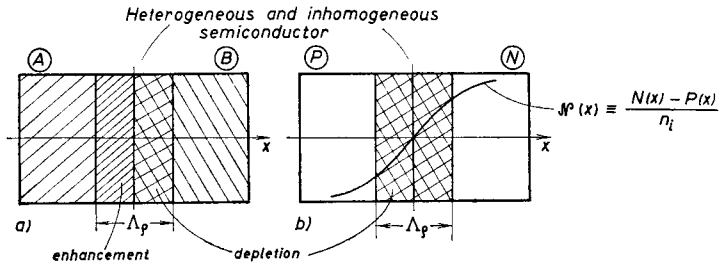

Fig. 1.1. — Heterogeneous ($a_A \neq a_B$, $m_A^* \neq m_B^*$, $\varepsilon_A \neq \varepsilon_B$, $W_{gA} \neq W_{gB}$, ...) and inhomogeneous ($a_P = a_N = a$, ..., $\varepsilon_P = \varepsilon_N = \varepsilon$ etc.) solid (junction). Λ_ρ — total SC-region width.

Data for some noval semiconductor materials T. 1.1

Mat.	W_g [eV]	μ_n [cm²/Vs]	P-N	Structures P-N-P	Ref.
GaAs	1,35	5 000	+	+	[17]
GaP	2,25	300	+	—	„
InSb	0,18	80 000	+	+	„
β-SiC	2,86	100	+	—	[a]

„+" — is applied: W_g energy gap; μ_n — electron mobility (ambient temperature), [a] Ref. [18] to [20] (concerning photoeffect).

Chapter 2 — SOME BASIC PROBLEMS OF P-N TRANSITION THEORY

Irrespective of the fact that the junction may be either inhomogeneous or heterogeneous even when no external field is applied, the space charge field exists in the vicinity of $x=0$**) and may have the intensity even in the range 10^3 to 10^4 V/cm. The condition of neutrality is, therefore, moved at a larger distance from the origin. Between *space charge region* (SC) and the *neutral region* (NR) there exists a narrow region of thickness Λ_{uP} or Λ_{uN} which may

*) The question of space charge region depletion when impurity centres are partially ionized is discussed in [44].

**) This point is cleary defined in Fig. 1.1 a, while for Fig. 1.1 b, $x=0$ for $\subset\!\!\!\supset \mathcal{N}=0$. In $x=0$, (where $\subset\!\!\!\supset \mathcal{N}=0$) which is called the point of *technological* (*physical*) transition, it is not necessary that $\rho=0$ and $(n-p)=0$. The point $x=x_{Kmax}$ where $\rho=0$ will be called *electric* transition point, and the point where $(n-p)=0$ will be called *inversion* point x_i (see also Fig. 2.3).

be designated as the *intermediate region* (IR) in which, as *in borderline case*, neither

(2.1) $$\varrho = 0 \quad \text{nor} \quad |en_i \mathcal{V}| \gg |(p-n)e|$$

holds.

In further discussion our attention will be focused to the phenomena occurring in the first (SC) and the last (IR) of the aforementioned regions.

2.1. — Transport relations and boundary conditions; quasi-neutrality criterion

a. — General transport relations

These relations were established long ago in their phenomenological form according to the band model (see for instance [8] to [15]; especially [9]), However, the transport coefficients in SC (and IR) regions (μ_s — mobility, $D_s = \frac{kT}{e}\mu_s$) and concentrations may depend explicitly on the field, which means that the recombination rates r_{en} and r_{ep} are also dependent upon $\boldsymbol{K}$. Hence it seams much better to emphasize these interrelationships by writing the continuity equation in the following form:

(2.2a) $$\frac{\partial n}{\partial t} = -r_{ne}(n, p, n_{ds}', p_{as}, n_{Rs''}) + \frac{1}{e}\nabla \boldsymbol{J}_n + g_n';$$

(2.2b) $$\left.\begin{array}{l} \cdots\cdots\cdots\cdots\cdots, \\ \dfrac{\partial n_R}{\partial t} = -\sum\limits_{j\neq k}\gamma_{kj} n_k (N_j - n_j) + \sum\limits_{j\neq k}\gamma_{jk}(N_k - n_k) n_j, \\ \cdots\cdots\cdots\cdots\cdots; \end{array}\right\} (A+D+R)\ eqns.$$

(2.2c) $$\frac{\partial p}{\partial t} = -r_{pe}(n, p, \ldots) - \frac{1}{e}\nabla \boldsymbol{J}_p + g_p',$$

where the subscript „ds'“ stands for s'-th donor, the subscript „as“ — for s-th acceptor and „$\mathcal{R}s''$“ for s''-th recombination level, and the subscript „k“ for every local level (A acceptors, D donor and R recombination levels). γ_{kj} are recombination coefficients (for electron transition $k \to j$), e—*absolute* value of electron charges, and $g_n'(\boldsymbol{r}, \boldsymbol{K}, \ldots)$ and $g_p'(\boldsymbol{r}, \boldsymbol{K}, \ldots)$ volume rate of *external* (not thermal — not intrinsic, say that of strong $\boldsymbol{K}'s$, light and similar) generation;

— for kinetic equations (for current densities $\boldsymbol{J}_m$):

(2.3a) $$\boldsymbol{J}_n(\boldsymbol{r}, t) = en(\boldsymbol{r}, \boldsymbol{K}) \cdot \mu_n(\boldsymbol{K}, \boldsymbol{r}) \cdot \boldsymbol{K}(\boldsymbol{r}, t) + eD_n(\boldsymbol{K}, \boldsymbol{r}) \cdot \nabla n,$$

(2.3b) $$\boldsymbol{J}_p(\boldsymbol{r}, t) = ep(\boldsymbol{r}, \boldsymbol{K}) \cdot \mu_p(\boldsymbol{K}, \boldsymbol{r}) \cdot \boldsymbol{K}(\boldsymbol{r}, t) - eD_p(\boldsymbol{K}, \boldsymbol{r}) \cdot \nabla p,$$

supposing that the current cannot flow through local levels ($\boldsymbol{J}_k = 0 = \nabla \boldsymbol{J}_k$, which is valid for „higher“ temperatures when the process of „hopping“ is neglected).

Hence, there are (1 + 1 + D + R + A) unknown „concentrations“ (of Felmi's quasi-levels $E_{fn}, \ldots$), two „unknown“ currents ($\boldsymbol{J}_n$ and $\boldsymbol{J}_p$) and the „unknown“ field or potential $\Upsilon(\boldsymbol{r}, t)$, with „known“ μ-s, D's, γ's, $N_{Rs''}$'s, r_{ne}'s and similar, it means that (2) and (3) result in (4 + A + D + R) equations i.e. regarding

the number of unknowns, one equation is missing, and therefore another equation must be taken*): Poisson's one

$$\nabla \boldsymbol{D} = \rho; \quad \boldsymbol{D} = \varepsilon_0 \varepsilon(\boldsymbol{r}) \cdot \boldsymbol{K}; \quad \boldsymbol{K} = -\nabla \Upsilon \equiv -\frac{1}{e} \nabla \varphi. \tag{2.4}$$

The system (2) to (4) was certainly well known even earlier (for instance [9], [15] and similar) but writing in the above *formulation one emphasizes* $\left(\text{transient case; } \frac{\partial n}{\partial t}, \frac{\partial p}{\partial t}, \ldots, \neq 0\right)$ in the general case:

— that the members $r_{ne} \equiv r_n - g_{tn}$ and r_{pe} *need not be* the same and that the writing of r_{ne} as $(n - n_0) \cdot \tau_n^{-1} \equiv \delta n \cdot \tau_n^{-1}$ (n_0 — equilibrium concentration) with constant lifetime τ_n is *not correct*; even only the direct recombination (c—v level) results in

$$\tau_n = (n - n_0)\,[\gamma_{cv}(np - n_i^2)]^{-1} = f(n, p);$$

— that $g_n' \neq g_p'$ and that it is better to *separate* the „external" generation effects (that is often left out, as for instance in [9] — Chapt. 3),

— that the dielectric constant (relative) can depend upon $\boldsymbol{r}$; $\varepsilon = \varepsilon(\boldsymbol{r})$ — for instance as in the junction in Fig. 1a: in heterojunctions.

— and that when applying ∇ to $\boldsymbol{J}_{n,p}$ from (2.3) one should pay attention to the fact that $\boldsymbol{D}_{n,p} = f(\boldsymbol{K})$, with $\boldsymbol{K} = \boldsymbol{K}(\boldsymbol{r}, \ldots)$.

b. — Accepted model

The system (2.2)—(2.4) cannot ce solved in its geneıal form, and the corresponding approximations**) based on definite physical conceptions should be carried out so that apart from the aforesaid, we shall suppose later on that:

— we have a steady state case:

$$\frac{\partial n}{\partial t}, \frac{\partial p}{\partial t}, \frac{\partial n_R}{\partial t} = 0; \tag{2.5}$$

(2.6a—c)
- — that $\nabla \varepsilon_{A,B} = 0$,
- — that all other „fields" ($\boldsymbol{B} = 0$ and $\nabla T = 0$ — isothermal conditions), except $\boldsymbol{K}$, do not effect,
- — that the state of carriers are not degenerated anywhere: $E_f(x)$ or $E_{fn,p}(x)$ does not intersect anywhere $E_c(x)$ or $E_v(x)$ (see also Fig. 2.1.),

— and that indirect recombination processes are sufficiently well represented by *one* $\mathcal{R}$ — level somewhere in the middle of the energy gap; hence (see for instance [15] — Chapt. 6):

$$r_{en} = r_{ep} = r_e = N_R \cdot \frac{np - B_{cR}\, B_{vR}}{(p + B_{vR})\,\gamma_p + (n + B_{cR})\,\gamma_n} \cdot \gamma_n \cdot \gamma_p = r_e(n, p),$$

$$\gamma_n \equiv \gamma_{cR}, \quad \gamma_p \equiv \gamma_{vR}; \quad B_{cR} \equiv B_c\, e^{-\tilde{W}_R} \approx B_{vR} \equiv B_v\, e^{\tilde{W}_R - \tilde{W}_g}, \tag{2.7}$$

$$B_{vR} \cdot B_{cR} = n_i^2; \quad B_{c,v} = 2\left(\frac{2\pi\, m_{gn,p} \cdot kT}{h^2}\right)^{3/2}.$$

*) The question of correctness of its application here as well as that of Einstein's relation ($D_s = \mu_s kT/e$) is a special problem which is not treated here.

**) Doubtless, the general relations (2.2)—(2.4) are of great significance then as well, because they form a correct and reliable starting point.

The choice of only one recombination level does not diminish the generality of phenomena and conclusions obtained.

The remark regarding notation: Normal subscripts as well as upper indexes are used according to the scheme:

(2.8 a) $$ {}^{z}A_{y}^{x}, $$

x — space and time notation:

$x=0$ means that $t=0$, $x=0$ or at some boundary ($x=x^{*}_{N,P}$ — Fig. 2.1);

$x=\infty$ means „far away from something";

y — defines the quantity A more closely (for instance $A_{n,p}$ — „what refers to the electrons" i.e. „holes", A_0 — equilibrium value);

z — approximation order.

(2.8 b) $$ \tilde{A} \text{ means } A/kT. $$

$\pm A_{b,c}$ (or $\pm A_{b,c}$) means A_b and $(-A_c)$. „$=\cdots=$" means after „longer" calculation.

See the list of symbols in the Appendix A.

c. — General boundary conditions

These conditions are different for the case in Fig. 1.1 a from those in Fig. 1.1 b, and therefore they are given separately.

1° — Regarding the *transition between two bodies* (A and B — Fig. 2.1 a) these conditions seem to be most completely formulated in [9] — Chapt. 3. The first condition is that (for smaller polarization) the resultant flow of the electrons $\left(-J_n^D\right)$ on the boundary $A-B$ (the layer C is „thin") must be equal to the difference $(J_{ABT}-J_{BAT})$ — of the nonequilibrium „thermal" currents ($A\to B$ and $B\to A$); the same is valid for the holes. Hence, the electron flow that may pass the barrier $\Phi_c\equiv E^{\circ}_{CA}-E^{\circ}_{CB}>0$ is (conventional direction):

$$ J_{BAT}=e\frac{m_{nB}^3}{4\pi^3\hbar^3}\int_{-\infty}^{+\infty}\int dv_y\,dv_z\int_{v_x=v_m}^{-\infty}(f_{0B}+f_{1nB})\,v_x\,dv_x=\cdots=J_{BAT^{\circ}}+\gamma_{nB}J^{\circ}_{nB}\,\mathrm{e}^{-\tilde{\Phi}_c}, $$

where:

$$ J_{BAT^{\circ}}=\frac{em_{nB}^3}{4\pi^3\hbar^3}\int_{-\infty}^{+\infty}\int dv_y\,dv_z\int_{v_m}^{-\infty}f_0\,v_x\,dv_x\approx\frac{4\pi\,em_{nB}}{h^3}\,(kT)^2\cdot\mathrm{e}^{\tilde{W}^{\circ}_{fB}-\tilde{\Phi}_c}, $$

(2.9)

$$ \gamma_{nB}J^{\circ}_{nA}\,\mathrm{e}^{-\tilde{\Phi}_c}\approx\frac{em_{nB}^3}{4\pi^3\hbar^3}\int_{-\infty}^{+\infty}\int dv_y\,dv_z\int_{v_m}^{-\infty}f_{1nB}\,v_x\,dv_x\equiv I_{nB}(v_m);\quad J^{\circ}_{nB}=I_{nB}(+\infty), $$

f_0 and f_1 — equilibrium and additional distribution function ($f=f_0+f_1$ — see for instance [15] — Chapt. 5 and 7).

$m_{n,p}$ are effective masses, $v_m=-\sqrt{\dfrac{2\,\Phi_c}{m_{nB}}}$.

The same expression is obtained for J_{ABT}, but in that case the index „B" should be changed to „A" and $\Phi_c = 0 = v_m$ is taken formally, while the electron flow A→B need not have to pass any barrier, and in addition $J^{\circ}_{nA} = -I_{nA}(+\infty)$. Finally according to (2.9):

$$J_n^D \equiv J^{\circ}_{nB} - J_{reB} = J^{\circ}_{nA} - J_{geA} = J_{BAT} - J_{ABT}$$

$$= \left(\frac{4\pi e m^{\circ}_{nB} k^2 T^2}{h^3} e^{\tilde{W}^{\circ}_{fB}} + \gamma_{nB} J^{\circ}_{nB}\right) e^{-\tilde{\Phi}_c} - \left(\frac{4\pi e m^{\circ}_{nA} k^2 T^2}{h^3} e^{\tilde{W}^{\circ}_{fA}} - \gamma_{nA} J^{\circ}_{nA}\right), \tag{2.10}$$

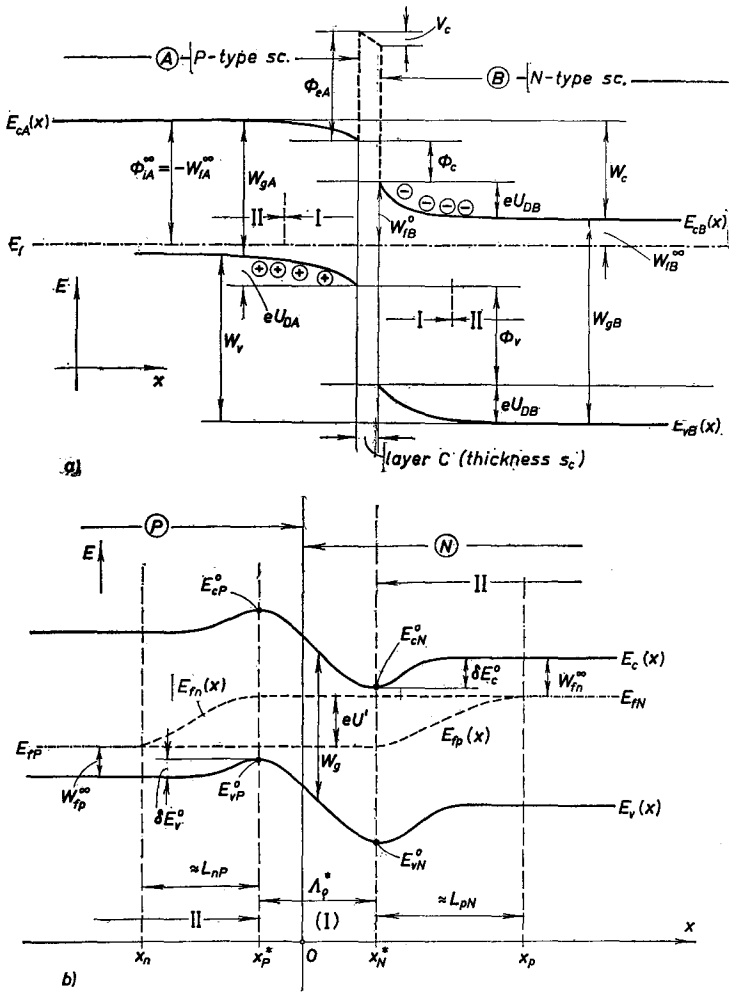

Fig. 2.1. — Energy situation in the vicinity of $x \approx 0$ for hetero (*a*) and P-N — transition (*b*).

a. — Two semiconductors contact (in equilibrium). U_D — diffusion potential Φ_i-s and Φ_e-s — internal and external work functions ($\Phi = \Phi_i + \Phi_e$), V_c — potential drop in the layer C.

b. — Band edges correction at direct polarization of N-P transition. U' — external voltage drop between x^*_P and x^*_N ($d\Upsilon/dx = 0 = K$ for that x); L_{nP} and L_{pN} — corresponding diffusion lengths ($L = \sqrt{D\tau}$); δE^0_c and δE^0_v — band edges changes at $x = x^*_N$ and x^*_P.

I — SC and IR regions, II — quasi-equilibrium (quasi-neutral) region (NR). „Relative" Fermi level $W_{fn} = E_{fn} - E_c < 0$. See the text for all information.

because there are also *recombination* currents J_{reB} on the boundary: $J_{reB} = J_{ncB-vA} = J_{nrB} - J_{ngB}$ ($J_{nrB} \equiv J_{nrcB-vA}$ and $J_{ngB} \equiv J_{ngvA-cB}$: transitions $c_B \rightleftarrows v_A$) and *generation* currents $J_{nvB-cA} \equiv J_{geA} = \cdots$ (due to the transitions $v_B \rightleftarrows c_A$), so that the whole J°_{nB} or J°_{nA} *does not pass* through the contact. From (Fig. 2.2a) the equation

$$W^{\infty}_{fB} - W^{\infty}_{fA} = \Phi_c + (eU_D + V_c), \quad e^{\tilde{W}^{\infty}_{fA}} = \frac{n^{\infty}_A h^3}{2(2\pi m^{\infty}_A kT)^{3/2}}, \quad e^{\tilde{W}^{\circ}_{fA}} = \cdots, \tag{2.11}$$

using (2.10) also, the *first boundary condition**) is obtained finally in the form

$$\frac{n^{\circ}_B}{n^{\infty}_B}\left(\frac{m^{\circ}_A\, m^{\infty}_B}{m^{\circ}_B\, m^{\infty}_A}\right)^{1/2} \cdot \frac{m^{\infty}_B}{m^{\infty}_A}\, e^{\tilde{eU}_D + V_c} = \frac{n^{\circ}_A}{n^{\infty}_A} + \sqrt{\frac{\pi m^{\circ}_A}{2kTe^2}}\left[\frac{J^{\circ}_{nB} - (J_{reB} + J_{geA})}{n^{\infty}_A} - \frac{J^{\circ}_{nB}}{n^{\infty}_B}\left(\frac{m^{\infty}_B}{m^{\infty}_A}\right)^{3/2} \cdot e^{\tilde{eU}_D + \tilde{V}_c}\right], \tag{2.12}$$

which is *more general* than that of (17. 19) in [9] and [15.2] — page 7.40, because it takes into account that $m_A \neq m_B$ and that the effective masses of electrons in the depth of any semiconductor are *different* from those on its surface ($m^{\circ}_{nA,B} \neq m^{\infty}_{nA,B}$) so that (2.12) is not transformed into (17.19) even if $m^{\circ}_A = m^{\circ}_B$, but only if $m^{\circ}_A = m^{\infty}_A = m^{\circ}_B = m^{\infty}_B = m_n$.

The *second* („electrostatic") condition is

$$K^{\circ}_{At} = K^{\circ}_{Bt} \quad \text{and} \quad D_{BN} - D_{AN} \equiv \varepsilon_B K_{BN} - \varepsilon_A K_{AN} = \sigma_{AB}, \tag{2.13a}$$

„t" — tangential, „N" — normal component, and σ_{AB} is here the surface charge of the layer C.

The *third* condition (which is very often forgotten) is:

$$E_{fB}(x^{*}_N) - E_{fA}(x^{*}_P) = eU' \quad \text{or} \quad E^{\infty}_{fB} - E^{\infty}_{fA} = eU', \tag{2.13b}$$

where the last condition is valid for „smaller" currents. The condition (2.13b) is valid also for the P-N transition (Fig. 2.2b).

2° — There is not any „discontinuity" for P-N transition at $x = 0$ (Fig. 2.1b), so that the boundary conditions are usually given (but not always, for instance in [22] they are given for $x = 0$) for transition „terminals" (the values for these points will be notified by the upper index „0") where the quasi-neutrality conditions prevail (about it see the point d). When the polarization is *direct*, p°_N and n°_P have not the equilibrium „somewhere" at $x_{N,P}$, because of injected minority carriers (holes from P- and the electron from N-region). Therefore the concentrations of the majority carriers must be increased in order that quasi-neutrality might be kept:

$$\begin{gathered} n^{\circ}_N - n^{\circ}_{oN} \approx p^{\circ}_N - p^{\circ}_{oN}, \quad n^{\circ}_{oN} + P^{\circ}_N = p^{\circ}_{oN} + N^{\circ}_N \;\therefore\; n^{\circ}_N \approx p^{\circ}_N + N; \\ p^{\circ}_P - p^{\circ}_{oP} \approx \cdots, \quad \cdots, \quad \therefore\; p^{\circ}_P \approx n^{\circ}_P + P; \\ N \equiv N^{\circ}_N - P^{\circ}_N \quad \text{and} \quad P \equiv P^{\circ}_P - N^{\circ}_P. \end{gathered} \tag{2.14}$$

*) Here we suppose that $\gamma_{nB} \approx \frac{1}{2} \approx \gamma_{nA}$, while treatment in the original version [21] is slightly more general.

As the currents $J_{nN;pP} \sim \nabla E_{fn;p}$ of the majority carriers in the vicinity of $x_{N;P}$ cannot change abruptly (see for instance [12]), $E_{fn,p}$'s remain almost constant here, and as $n_N^\circ > n_{oN}^\circ$ and $p_P^\circ > p_{oP}^\circ$, the band edges $E_{c,v}$ must „correct" this situation, approaching to E_f's (at $E_{cN;vP}^\circ$). As it was shown in [21] (see also Fig. 2.1b), on the basis of (2.14) and

$$p_N^\circ n_N^\circ = n_i^2 e^{e\tilde{U}'} = p_P^\circ n_P^\circ = pn \quad \text{for} \quad x_P^* \leqslant x \leqslant x_N^* \tag{2.15}$$

the following *conditions* can be obtained:

$$\boxed{(2.16a)} \qquad p_N^\circ = N \frac{\sqrt{1+4\frac{P}{N}\chi}-1}{2}, \quad n_P^\circ = P \frac{\sqrt{1+4\frac{N}{P}\chi}-1}{2}; \quad \chi = \frac{n_i^2}{PN} e^{e\tilde{U}'};$$

correction — the displacement of band edges is

$$\boxed{(2.16b)} \qquad \delta \tilde{E}_c^\circ = \ln\left[\frac{1}{2}\left(\sqrt{1+4\frac{P}{N}\chi}+1\right)\right], \quad \delta \tilde{E}_v^\circ = \ln\left[\frac{1}{2}\left(\sqrt{1+4\frac{N}{P}\chi}+1\right)\right],$$

$$\delta W_f^\circ = \delta E_c^\circ + \delta E_v^\circ = kT \ln \frac{n_N^\circ \cdot p_P^\circ}{N \cdot P}, \quad \left(e\tilde{U}_D \equiv \ln \frac{PN}{n_i^2}\right),$$

that is valid practically for all injection levels (when $\chi \leqslant 1$ or $U_D \geqslant U'$) and for the reverse polarization.

The analysis of (16a) can bring us to many useful conclusions. Thus, for instance, for $\chi = 1$, for symmetrical ($P = N$) and for extremely asymmetrical ($P \gg N$) transition

$$p_N^\circ = N \frac{\sqrt{5}-1}{2} = 0.618\, N = n_P^\circ < N = P \quad \text{and}$$

$$p_N^\circ \approx \sqrt{N \cdot P}, \quad n_P^\circ \approx N \equiv n_{oN}^\circ \ll P \tag{2.16c}$$

is obtained respectively, that means that the injected concentration never *exceeds* (if $\chi \leqslant 1$) their equilibrium values ($P = p_{oP}^\circ$ and $N = n_{oN}^\circ$) in their region. This conclusion seems „*a priori*" logical, but if the correction is not taken into account properly, starting from (2.14) one has

$$n_P^\circ = \frac{\chi N + \chi^2 P}{1-\chi^2} \quad \text{and} \quad p_N^\circ = \frac{\chi P + \chi^2 N}{1-\chi^2} \tag{2.16d}$$

instead of (2.16a) like for instance in [12—153]. Consequently, an absurd conclusion that n_P° and $p_N^\circ \to \infty$ when $\chi = 1$, will follow.

Apart from that, such presentation enables the precise definition of $x_{N,P}^*$; here is $K = 0$ *exactly* (and it changes the sign), while without it (Shockley's conception [24]) for $x = x_{N,P}$ (without the asterisk) we have $K \approx 0$ somehow „asymptotic". Thus if the transition is abrupt or linear (in total depletion approximation — [24]) we obtain:

$$x_N^* = -\frac{P}{N} x_P^* = \left[\frac{2\varepsilon\varepsilon_0 U^+ P}{eN(N+P)}\right]^{1/2} = \frac{P}{P+N}\Lambda_\rho^*,$$

$$x_N^* = -x_P^* = \frac{\Lambda_\rho^*}{2} = \left(\frac{3\varepsilon\varepsilon_0 U^+}{2\varrho_0 e}\right)^{1/3}, \tag{2.17b}$$

respectively, but with

(2.17c) $$eU^+ \equiv 2\alpha^* kT = eU^* + \delta W_f^\circ, \quad U^* = U_D - U'$$

and $\mathcal{S}_0 = n_i \frac{d\mathcal{N}}{dx} = C^{te}$ — linear transition slope.

d. On the quasi-neutrality criterion

Putting $\mathcal{N}_R = 0$ in (1.1) with

$$n = n_0 + \delta n = B_c\, e^{\tilde{W}_{fn}} = \cdots = n_i\, e^{\tilde{\varphi} - \tilde{\mu}_{fn}}, \quad p = p_0 + \delta p = n_i\, e^{\tilde{\mu}_{fp} - \tilde{\varphi}};$$

(2.18) $$\varphi = E_f - E_i \equiv E_f - \left[\frac{E_c(r) + E_v(r)}{2} + \frac{kT}{2} \ln \frac{B_v}{B_c}\right] = \psi + \frac{\mu_{fn} + \mu_{fp}}{2},$$

$$\mu_{fp,n} = E_f - E_{fp,n}, \quad \mu_{fp} - \mu_{fn} = 2\beta,$$

under the condition that $\beta = eU'/2 = C^{te}$ (which holds good for direct polarization in $x_N^* > x > x_P^*$ — see Fig. 2.1 b), (2.4) becomes ([21])

(2.19) $$\Delta\tilde{\psi} = \frac{1}{\Lambda_{Di}^2}\left[e^{e\tilde{U}'/2} \cdot 2 \sinh \tilde{\psi} - \mathcal{N}(r)\right], \quad \Lambda_{Di}^2 = \frac{\varepsilon\varepsilon_0 kT}{e^2 n_i},$$

so that for $U = 0 = U'$ ($\varphi = \psi$) the *quasi-neutrality condition* is the following:

(2.20a) $$2 \sinh \tilde{\varphi} - \mathcal{N}(r) \equiv -\rho/en_i \approx 0 \quad \text{or} \quad \left|\frac{\rho}{en_i}\right| \equiv |2 \sinh \tilde{\varphi} - \mathcal{N}| \ll |\mathcal{N}|,$$

which is well known (see for instance [13] where, however, it is not emphasized that absolute values should be compared), but determining $\Delta\tilde{\varphi}$ from (2.20a) it is assumed (for instance in [13]) that from

(2.20b) $$\rho \approx 0 \quad \text{it results also} \quad \nabla\rho \approx 0 \quad \text{and} \quad \Delta\rho \approx 0,$$

while *without* this assumption *quasi-neutrality criterion* will be obtained in the form which is *more general:*

(2.20c) $$\left|\Delta\left(\mathcal{N} - \frac{\rho}{en_i}\right) - \frac{\mathcal{N}}{4 + \mathcal{N}^2}\left[\nabla\left(\mathcal{N} - \frac{\rho}{en_i}\right)\right]^2\right| \ll \frac{|\mathcal{N}|}{\Lambda_{Di}^2}\sqrt{4 + \mathcal{N}^2}.$$

That (formal) taking of $\rho = 0$ in (2.20c), as for instance in [13], is not always correct, it can be seen from the example of abrupt transition (Fig. 2.2). Actually, as $\nabla\mathcal{N} = 0 = \Delta\mathcal{N}$ for $x > 0$ for the abrupt transition, with $\rho = 0$ eqn. (2.20c) would be

(2.20c') $$0 \ll \mathcal{N}\sqrt{4 + \mathcal{N}^2}, \qquad \mathcal{N} = \mathcal{N}_N,$$

which is always fulfilled, and the result of which (evidently incorrect) would be that the *quasi-neutrality conditions* are fulfilled everywhere in this region. However, it is obvious from Fig. 2.2. that neither $\Delta_x\rho$ nor $\nabla_x\rho$ are equal to zero, so that the left hand side in (2.20c) which is actually proportional to ρ, *never* equals zero, and it decreases abruptly starting from centrain $x_{\rho N}$, which

will be defined in Section 2.2. Similar statement might be done for the graded transition.

Eqn. (2.20c) is also valid for small reverse polarizations and for *almost all* direct polarizations, provided that $(4\beta^2)$, $\beta^2 = \exp e\tilde{U}'$, should be written everywhere instead of 4. Thus (2.20c) is fulfilled more satisfactorily for direct polarization, and less satisfactorily for reverse polarization. See also the eqn. (2.28c).

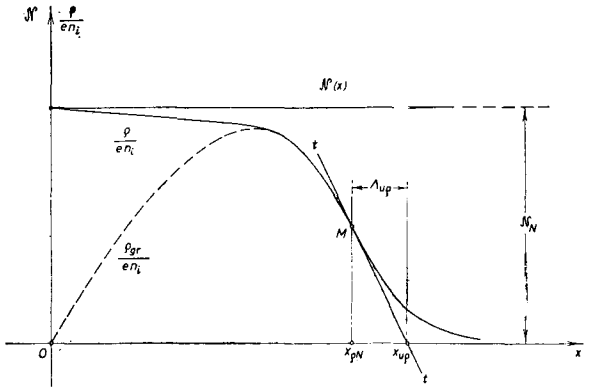

Fig. 2.2. — Example of incorrect application of (2.20c) with $\rho = 0$ for the abrupt transition (N-region). ρ_{gr} — for graded transition; M — inflection point ($x_{\rho N} \equiv x_M$); tt — tangent in M.

2.2. New approximate solutions of Poisson's equation (small polarizations)

a. — Solution conception

Poisson's equation (2.19) in its general form cannot be solved even not for the equilibrium case, which is assumed here also ($\psi = \varphi$, $U = 0 = U'$). Nevertheless, in the intermediate region (somewhere at $x_{\rho,N}$ — Fig. 2.2) the potential φ lacks „very little" to reach its *asymptotic* values:

$$\tilde{\varphi}_{N,P} = \sinh^{-1}\frac{\tilde{\mathcal{N}}}{2}, \quad \rho \to 0, \quad (x \to \pm\infty), \tag{2.21a}$$

although $(n-p) \sim \sinh\tilde{\varphi}$ changes very much here, and the conditions (2.20c) *need not* be fulfilled. Under the *assumption* that here (or anywhere else) one has

$$\frac{d}{dx} 2\sinh\tilde{\varphi} \gg \frac{d\tilde{\mathcal{N}}}{dx} \text{ and with } z = \frac{d\tilde{\varphi}}{dx}, \quad \left(\tilde{\varphi}'' = z\frac{dz}{d\tilde{\varphi}}\right), \tag{2.21b}$$

we have obtained in [21]:

$$\boxed{(2.21c)} \quad \frac{2x}{\Lambda_{Di}} + C = \int \frac{d\tilde{\varphi}}{\left[Z_u^2 + \cosh\tilde{\varphi} - \cosh\tilde{\varphi}_u + \tilde{\varphi}_u \frac{\tilde{\mathcal{N}}_u}{2} - \frac{\tilde{\mathcal{N}}(x)}{2}\tilde{\varphi}\right]^{1/2}} \approx$$

$$\approx \int \frac{d\tilde{\varphi}}{\left[Z_u^2 + \left(\frac{\tilde{\mathcal{N}}_u}{2} - \frac{\tilde{\mathcal{N}}}{2}\right)\tilde{\varphi} + \cosh\tilde{\varphi}_u \frac{(\tilde{\varphi}-\tilde{\varphi}_u)^2}{2!} + \sinh\tilde{\varphi}_u \frac{(\tilde{\varphi}-\tilde{\varphi}_u)^3}{3!} + \cdots\right]^{1/2}}$$

where the last expression represents the expansion of $\cosh\tilde{\varphi}$ (near $\tilde{\varphi}_u$) in the series and where:

$$(2.21\,\mathrm{d}) \qquad \tilde{\varphi}_u=\tilde{\varphi}(x_u),\quad Z_u=\frac{z_u\Lambda_{Di}}{2},\quad z_u=z(x_u)=\frac{\left|\nabla\mathscr{N}_u-\nabla\frac{\rho_u}{en_i}\right|}{2\cosh\tilde{\varphi}_u}\approx\frac{|\nabla\mathscr{N}_u|}{4+\mathscr{N}_u^2},$$

provided that the quantities with subscript „u" are constant, and $x_u=\mathrm{const}$ or $x_u\to\infty$.

The actual solution choice consists of retaining the largest number of terms in expansion but the solutions must be given with elementary functions.

b. — Abrupt transition

The assumption (2.21 b) for the abrupt transition is quite *correct* ($\mathscr{N}=\mathscr{N}_u=n_i N=\mathrm{const}$) and Z_u is very near to zero (this holds only approximately for graded transitions). By taking $Z_u=0$ and $\mathscr{N}_u=\mathscr{N}$, and by retaining the terms to the fourth power*) inclusively, in (2.21 c),

$$(2.22) \qquad \tilde{\alpha}\equiv\tilde{\varphi}_u-\tilde{\varphi}=\frac{12C_2'\,e^{x/\Lambda'_{Di}}}{(C_2'\,e^{x/\Lambda'_{Di}}+\tanh\tilde{\varphi}_u)^2-3}=$$

$$=\frac{6}{\sqrt{3-\tanh^2\tilde{\varphi}}\cdot\sinh\left(\frac{x}{\Lambda'_{Di}}+C_2''\right)+\tanh\tilde{\varphi}_u},\qquad \Lambda'_{Di}=\frac{\Lambda_{Di}}{\sqrt{2\cosh\tilde{\varphi}_u}},$$

is obtained [21]. It is obvious that $\tilde{\alpha}(x)$ decreases more slowly than $\exp(-x/\Lambda'_{Di})$ when $x\approx x_{\rho N}$ (for larger $\tilde{\alpha}$'s) and that $\tilde{\alpha}\sim\exp\left(-\frac{x}{\Delta'_{Di}}\right)$ if $x\to\infty$ (smaller $\tilde{\alpha}$'s).

The expression (2.22) is more general than any of those given up to now (for instance those in [23] and [25]) and it might be reduced to the latter. For instance, Shockley (from the expression 2.13 in [24]) and Aigrain ([12] — page 138) obtain

$$(2.23\,\mathrm{a\ and\ b}) \qquad \frac{d^2\tilde{\alpha}}{dx^2}=\frac{\tilde{\alpha}}{(\Lambda'_{Di})^2}\quad\therefore\quad \tilde{\alpha}\sim e^{-x/\Lambda'_{Di}}\ \text{and}\ \tilde{\alpha}\sim e^{-x/\Lambda_{DN}},\quad \Lambda_{DN}=\left(\frac{\varepsilon\varepsilon_0 kT}{e^2N}\right)^{1/2},$$

respectively, for the intermediate region when α is *very small* in the „space charge" approximation (total depletion for $x\leqslant x_N$); (2.23 a) can be obtained from (2.21 c) neglecting all the terms except the square one, or directly from (2.22) for smaller $\tilde{\alpha}$; eqn. (2.23 b) is more narrow than (2.23 a) because it is supposed that $N\gg n_i$ ($N\approx n_i\exp\tilde{\varphi}_u$) i. e. that the condition $\tilde{\varphi}_u\gg 1$ must be satisfied; (2.22) may be shown to be valid for $\tilde{\alpha}$'s even when they are *larger* than 1.

*) Only the second power is retained in [23], so that (2.22) is much more correct. If higher powers (from the forth on) were retained in (2.21 c) as well, the solutions would be even more correct, but not expressed within *elementary functions*.

Another advantage of the relation (2.22) is in the fact that all its solutions are *analytical* and therefore they make analysis easy, what is not possible with the solutions obtained by computers (as for instance in [26]), although they are more accurate. Apart from that, the expressions of the type (2.23) are obtained under additional assumption that besides

(2.23c) $$\tilde{\alpha} \ll \tilde{\varphi}_u \text{ one must have } \tilde{\alpha} \equiv \frac{\alpha}{kT} \ll 1,$$

while for (22) the last condition is *not required.*

c. — Application to asymmetrical transition

Let us divide the whole region into four parts (Fig. 2.3.):

—(I), $x_P \leqslant x \leqslant 0$, where (2.22) is valid with

(2.24a) $$\tilde{\varphi}_u - \tilde{\varphi}_P = \sinh^{-1}\left(\frac{-P}{2n_i}\right) \text{ and } \alpha_P = \varphi^\circ - \varphi_P \approx 0$$

because for $P \gg N$ the whole potential drop is in the N-region;

—(IV), $x_i \leqslant x \leqslant x_N$, where also (2.22) is valid with

(2.24b) $$\tilde{\varphi}_u - \tilde{\varphi}_N = \sinh^{-1}\frac{N}{2n_i} = \tilde{\alpha}_N,$$

or eqn. in TB.a*) if the origin is transferred to x_i;

— (III), near $x \approx x_i$; only in this part carrier charge (according to Shockley — parabolical dependence of $\tilde{\varphi}$) *can be neglected,* and

— (II), $0 \leqslant x \approx x_i$ or $0 > \varphi^\circ \leqslant \varphi \leqq \varphi_c \approx 0$, where (2.22) *cannot* be used, because for $\tilde{\varphi} < \tilde{\varphi}_c$ it is not valid $\left(\tilde{\varphi}_c \approx \tilde{\varphi}_N - \frac{6}{\tanh \tilde{\varphi}_N}\right.$ — the lower boundary for approximation $\tilde{\varphi}_{IV}$, see also Fig. 2.3$\left.\right)$.

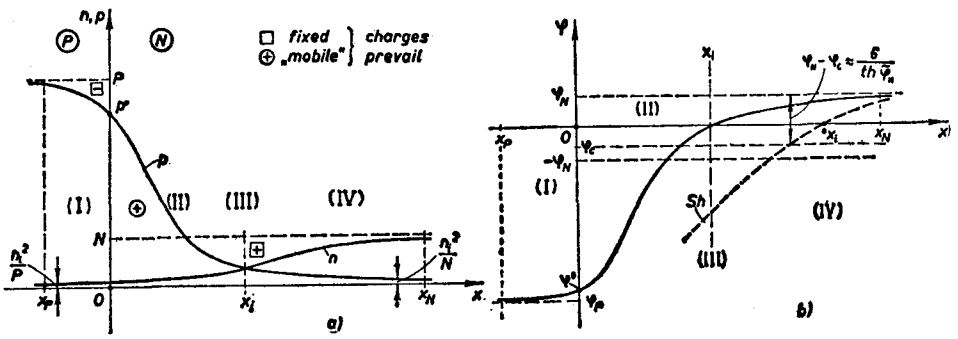

Fig. 2.3. — Dependences $p(x)$, $n(x)$—(*a*) and $\varphi(x)$—(*b*) for asymmetrical ($P \gg N$) abrupt transition (in equilibrium $U = O = U'$).

*) In order to clarify the text, the list of numerous clumsier expressions is given in the Appendix B, see the table TB.a/TB.b and further on; the potentials for some particular regions bear the subscripts I to IV.

It is obvious that the solution for P-region (Fig. 2.3) is not interesting and that separate solutions for (III) *may be left out*, because (2.22) is more correct than other solutions and may be used for larger $|\tilde{\alpha}|$'s, so that $\tilde{\varphi}_{IV}$ and $\tilde{\varphi}_{II}$ *overlap with* $\tilde{\varphi}_{III}$ (or mutually) in (III).

Thence, from the point of view of our solution (2.22) the only problem is the region (II), where the following possibilities based on (2.21 c) are examined.

1° — Since at $x=0$ we have

$$p=n_i\,e^{-\tilde{\varphi}_{II}}\gg n=n_i\,e^{\tilde{\varphi}_{II}}\quad\therefore\quad \cosh\tilde{\varphi}_{II}\approx e^{-\tilde{\varphi}_{II}}/2\gg\frac{N}{2n_i}\,|\tilde{\varphi}_{II}|, \tag{2.24 c}$$

instead of $\frac{\mathcal{N}}{2}\tilde{\varphi}$ the constant ν is introduced ($\nu=0$ is taken in [23], while here $0\leqslant\nu\leqslant 1$, what is more „flexible"), so that the solution is as given in TB.b; the subscript „e" is to remind us to the „exponential" dependence $\exp\tilde{\varphi}_{II}$ instead of $\cosh\tilde{\varphi}$.

2° — That $\cosh\tilde{\varphi}$ expands again into a series, is the second possibility, but near some „nonasymptotic" value $\tilde{\varphi}_1$ ($0>\tilde{\varphi}^\circ\leqslant\tilde{\varphi}_1\leqslant\tilde{\varphi}_c$). The solution for $\tilde{\varphi}_{II}$ with

$$\tilde{\varphi}_1=-\tilde{\varphi}_N<0 \tag{2.24 d}$$

when it stops in its expansion at a square term (in order to avoid special functions, since now there is a constant term) it is shown in TB.b down, provided that x° is determined from the condition $\varphi_{II}=\varphi^\circ$ for $x=0$ with $\tilde{\varphi}^\circ$ and x° given in TB.d.

3° — Near $x=x_i$ where $\tilde{\varphi}\approx 0$ and for large $|\tilde{\varphi}|$'s i. e. everywhere where

$$\cosh\tilde{\varphi}\gg\left|\frac{\mathcal{N}}{2}\tilde{\varphi}\right|, \tag{2.24 e}$$

eqn. (2.21 c) is reduced to the *elliptic integral* of the first kind: no series expansion is required (when $Z_u^2\approx 0$) and the solution for $\tilde{\varphi}_{el}$ is like that in TB.c.

Inversion points x_i can also be determined on the basis of the above mentioned, from the *condition* that $x=x_i$ for $\tilde{\varphi}=0$:

(II). — If $\tilde{\varphi}_c\geqslant 0$ (larger $\tilde{\varphi}_N$'s and N's) then $x_i\equiv x_{iII}$ is determined directly from the expression for $\tilde{\varphi}_{II}$ in TB.b — below, because at $\tilde{\varphi}\approx 0$ the $\tilde{\varphi}_{II}$ should not be the worse approximation than $\tilde{\varphi}_{IV}$. Thus the expression given in TB.d is obtained for x_{iII}.

(IV). — If, however, $\tilde{\varphi}_c<0$ (smaller $\tilde{\varphi}_N$'s and N-s, the case of silicon transition in [23]), $\tilde{\varphi}_{II}$ and $\tilde{\varphi}_{IV}$ must first be „joined" (determining C_2' in TB.a), then $x_i\equiv x_{iIV}$ is obtained from $\tilde{\varphi}_{IV}\approx 0$. From the condition $\tilde{\varphi}_{II}(x_c)=\tilde{\varphi}_c\equiv\tilde{\varphi}_{IV}(x_c)$ with $\tilde{\varphi}_c$ according to TB.e (see also Fig. 2.3.) x_c and *finally* x_{iIV} are obtained, as in TB.d. — below.

(IIe) and (IVe). — If we use the expression for $\tilde{\varphi}_{IIe}$ (TB.b — below, see also the paragraph 1°) instead of $\tilde{\varphi}_{II}$ by means of the same procedure

— x_{iIIe} as in TB.e — above is obtained for $\tilde{\varphi}_c\geqslant 0$, and

— x_{iIVe} as in TB.e — below is obtained for $\tilde{\varphi}_c<0$.

T 2.1. — *Results of calculation for* x_i*'s according to TB.a—e for cases A, B and C*

Results for / Quantity	A $\frac{N}{n_i}=0.26$ $\frac{P}{n_i}=2.6\cdot10^4$	B $\frac{N}{n_i}=5.78$ $\frac{P}{n_i}=5.78\cdot10^5$	C $\frac{N}{n_i}=1.3\cdot10^4$ $\frac{P}{n_i}=1.3\cdot10^8$	Calculated according to
$\tilde{\varphi}_N$	0.13	1.78	9.47	(2.21a)
$\sinh\tilde{\varphi}_N=\eta_N$	0.13	2.89	6500	
$\tanh\tilde{\varphi}_N$	0.129	0.95	≈1	
$\cosh\tilde{\varphi}_N$	1.0085	3.05	6500	
$\cosh\tilde{\varphi}_P\cdot10^{-4}$	1.3	28.9	6500	
$-\tilde{\varphi}_P$	10.17	13.27	18.68	(2.21a)
$\sqrt{2\coh\tilde{\varphi}_N}$	1.42	2.47	115	
$100\,M$	1.14	88	291	TB.b
$-\tilde{\varphi}^\circ$	9.17	12.27	17.68	TB.d
x°/Λ'_{Di}	6.7	2.64	1.43	TB.d
x°/Λ_{Di}	4.76	1.07	1.16×10^{-2}	(2.22)
$x_{i\,II}/\Lambda'_{Di}$	4.32	2.57	2.4	TB.d
$x_{i\,II}/\Lambda_{Di}$	3.02	1.04	$\mathbf{2.18\times10^{-2}}$	(2.22)
$x_{i\,IV}/\Lambda_{Di}$	1.70	0.92	$\mathbf{2.14\times10^{-2}}$	TB.d
$100\,M_N$*)	6.465	71	206	TB.d
C_N*)	8.4×10^3	3.1	5.4×10^4	TB.b
$\exp\frac{\tilde{\varphi}^\circ}{2}$	1/96	1/470	1/6750	
$\tilde{\varphi}_C$	−46	−4.52	3.47	TB.e
$100\,x_e^\circ/\Lambda_{Di}$	−1.48	−0.31	−0.021	TB.e
$x_{i\,II\,e}/\Lambda_{Di}$	1.4	0.93	2.72×10^{-2}	TB.e
$x_{i\,IV\,e}/\Lambda_{Di}$	**2.8**	**0.68**	3.25×10^{-2}	TB.e
x_{icl}/Λ_{Di}	1.3	1	3.68×10^{-2}	TB.c
$^\circ x_i/\Lambda_{Di}$	7.6	1.49	2.77×10^{-2}	TB.e
x_i/Λ_{Di}	2.35	0.77	2.17×10^{-2}	[23]

*) $\nu=1$; the values almost do not depend upon ν for the case *C*.

The last expression in TB.e for ${}^{\circ}x_i$ represents the zero approximation (total exhaustion of carriers i.e. the so-called ,,space charge" approximation — of Shockley [24]), which *always* gives much *larger* values (see also the curve Sh in the Fig. 2.3.). This approximation is called Sh-approximation in the further text.

According to the expressions in TB.a — TB.e (they are deduced in [21]), the corresponding x_i's can be calculated for the given values N and P, what is already done in T 2.1. for the three pairs (A, B and C) of values (N, P) taken from [23], where x_i is calculated by computer (the last column in T 2.1). The results for some auxilliary values are also given in T 2.1. Our values for x_i's are compared with the correct ones (according to [23]), and analysed in Chapter 4.

Finally, the dependences $\varphi_r\left(\frac{x}{\Lambda_{Di}}\right)$, $(r = \mathrm{II}, \mathrm{II_e}, \mathrm{IV};\ el)$, can be calculated according to the expressions in the same tables, what will be done in Chapter 4.

d. — Intermediate region width ($\Lambda_{u\varrho}$ in Fig. 2.2) is not ,,sharply" defined, but on the basis of (2.22) and of the definition in Fig. 2.2. (which is valid for all kinds of transitions because ϱ always has the inflection point, while the abrupt transition has not, for $x \neq 0$, a maximum of ϱ), it can be written down as ([21], see also eqns. 2.23):

$$\boxed{(2.25\,\mathrm{a})} \qquad \Lambda_{u\varrho} = \frac{2\sinh\tilde{\varphi}(x_M) - \mathcal{N}(x_M)}{\mathcal{N}'(x_M) - 2\cosh\tilde{\varphi}(x_M)\cdot\tilde{\varphi}'(x_M)} = \Lambda_{Di}^2 \frac{\tilde{\varphi}'(x_M)\cdot\tanh\tilde{\varphi}(x_M)}{2\cosh\tilde{\varphi}(x_M)},$$

where the last expression is valid for the *abrupt* transition. It is obvious that $\Lambda_{u\varrho}$ *decreases* with the increase of T and that it is ,,*small*" because $\tilde{\varphi}'(x_M)$ is small, and apart from that $\Phi \equiv \frac{\tanh\tilde{\varphi}}{2\cosh\tilde{\varphi}} \leqslant 0.25$ (maximum of Φ is for $\sinh^2\varphi = 1$). If (2.17 b) is applied to the abrupt and very asymmetrical transition ($P \gg N$), for equilibrium $U = 0 = U'$ we get the following

$$\boxed{(2.25\,\mathrm{b})} \qquad \frac{\Lambda_\varrho}{\Lambda_{u\varrho N}} \geqslant \sqrt{\frac{16\,n_i\cdot(-\tilde{\varphi}_P)}{N[\tilde{\varphi}_N - \tilde{\varphi}(x_M)]^2\cosh\tilde{\varphi}_N}} \approx \frac{8\frac{n_i}{N}}{[\tilde{\varphi}_N - \tilde{\varphi}(x_M)]}\sqrt{\frac{(-\tilde{\varphi}_P)}{2}},$$

where the last expression is valid if $2\cosh\tilde{\varphi}_N \approx \frac{N}{n_i} \geq 5$. As $[\tilde{\varphi}_N - \tilde{\varphi}(x_M)]$ is of the order of unity and smaller, the *conclusion* is that Λ_ϱ ($x_{\varrho N}$ — in this case) may be both

— *larger* (higher T, direct polarization),

— and *smaller* (lower T, reverse polarization) than $\Lambda_{u\varrho N}$; this ratio is of the order of unity for $U = 0$ and ,,medium" T's.

2.3. — The question of space charge region existence

It is evident that the abrupt P-N transition will *always* have the space charge region (depleted of carriers). Nevertheless, the graded transition may also be quasi-neutral. The criterion for the transition slope ($\mathcal{S}_0$), its dependence upon *temperature* (T) and applied *voltage* are examined here. In the end, the recombination process influence is given.

a. — The required transition slope ($U=0$)

Let the transition be *linear* at the place where $x=0$:

$$(N-P)\equiv n_i \mathcal{N}(x)=\mathcal{S}_0 x, \quad \left(\mathcal{S}_0=n_i|\nabla \mathcal{N}|\equiv\frac{n_i}{\Lambda_0}\right) \tag{2.26a}$$

and symmetrical: $\rho(0)=0$, since then $n(0)=n_i=p(0)$. Evidently, the space charge will not exist if the quasi-neutrality condition (2.20c) is satisfied. This condition together with (2.26a) gives

$$\boxed{(2.26b)} \qquad \left(\frac{\mathcal{S}_0}{n_i}\right)^2\equiv\left(\frac{1}{\Lambda_0}\right)^2\ll\frac{(4+\mathcal{N}^2)^{3/2}}{\Lambda_{Di}{}^2}=\frac{8}{\Lambda_{Di}^2}$$

$$\therefore \left(\frac{\Lambda_0}{\Lambda_{Di}}\right)^2\gg\frac{1}{8} \quad \text{or} \quad \Lambda_0\gg\Lambda_{Di} \quad \text{and} \quad \mathcal{S}_0\ll\frac{n_i}{\Lambda_{Di}},$$

because one has $|\nabla\mathcal{N}|\gg|\nabla\rho|/en_i$, in the quasi-neutral case, and $N=0$ for $x=0$.

Evidently, (2.26b) does not give us the correct answer to the criterion for the space charge region existence, but if (2.26b) is satisfied — SC certainly *does not exist*, and if:

$$\Lambda_0\ll\Lambda_{Di} \quad \text{or} \quad \mathcal{S}_0\gg\frac{n_i}{\Lambda_{Di}} \tag{2.26c}$$

the SC-region will certainly exist.

We must mention that the condition (2.26b) for quasi-neutrality is usually written as $\Lambda_0\gg\Lambda_{Di}$ (for instance in [9] — page 242), but we have proved here that the condition $\Lambda_0\geqslant\Lambda_{Di}$ is sufficient; apart from that, the reason why the *transition slope* $\mathcal{S}_0$ must be compared just with Debye length i.e. with $n_i/\Lambda_{Di}\equiv\mathcal{S}_i$, is not often shown, but that dependence is clearly seen from (2.26b).

The space charge region or rather the region of the *depletion* of carriers will actually exist when the following condition is satisfied:

$$2|\sinh[\tilde{\varphi}(r)]|\ll|\mathcal{N}(r)| \quad \text{or} \quad \Delta\tilde{\varphi}\approx-\frac{\mathcal{N}}{\Lambda_{Di}^2}; \tag{2.26d}$$

it means that it is necessary to solve the complete Poisson's equation (2.19) and then to see whether (2.26d) is satisfied. For our rough estimation, we shall solve (2.26d) — the right hand side, with $\mathcal{N}$ according to (2.26a), under the condition (Shockley) that SC of carriers may be neglected up to $x=x_N$ when, „at once" the neutral region (NR) begins:

$$\varphi=\frac{x_N{}^2 x}{2\Lambda_0\Lambda_{Di}{}^2}-\frac{x^3}{6\Lambda_0\Lambda_{Di}^2}; \tag{2.26e}$$

in order to determine x_N we have the equations:

$$\sinh\tilde{\varphi}(x_N)=\sinh\frac{x_N{}^3}{3\Lambda_0\Lambda_{Di}^2}=\frac{x_N}{2\Lambda_0} \quad \text{for} \quad x_N<x_2 \tag{2.26f}$$

or

$$\sinh\frac{\Lambda_0{}^2 N_N{}^3}{3\Lambda_{Di}^2 n_i{}^3}=\frac{N_N}{2n_i}, \qquad \text{for} \quad x_N=x_2\equiv\Lambda_0\frac{N_N}{n_i},$$

where x_2 represents the distance after which ($x\geqslant x_2$) impurity density remains constant ($n_i\mathcal{N}=N_N$ — the final concentration).

Let $a=(3\Lambda_0\Lambda_{Di}^2)^{-1}\sim 1/\Lambda_0$, then

$$\left(\frac{d}{da}\sinh ax_N{}^3\right):\left(\frac{d}{da}Aax_N\right)$$

will always be bigger than one ($A\equiv 3\Lambda_{Di}{}^2/2$) i.e. the bigger the slope, $\sinh ax_N{}^3$ will increase more than Aax_N and their intersection according to (2.26f) takes place for smaller x_N. It means that with the *increase* of σ_0, x_N will *decrease*, but the *ratio* $x_N/\Lambda_0\sim\sigma_0 x_N$ *increases*. Therefore if the final N_N is given, it is *easier* to achieve x_N to be larger than x_2. It means that the depleted layer „penetrates" into the region $N=$const. more easily. *The least transition slope required* for this penetration will be:

(2.26g)
$$\frac{\sigma_{0\,min}}{n_i}=\left(\frac{1}{\Lambda_0}\right)_{min}=\frac{1}{\Lambda_{Di}}\sqrt{\frac{(N_N/n_i)^3}{3\sinh^{-1}\frac{N_N}{2n_i}}}\equiv y\approx\frac{1}{\Lambda_{Di}}\sqrt{\frac{(N_N/n_i)^3}{\ln\left(\frac{N_N}{n_i}\right)^3}},$$

where the final expression is valid for semiconductors of one type dominating $\left(\frac{N_N}{n_i}\gg 1\right)$.

It can be easily proved that, at a certain definite temperature, (n_i, Λ_{Di}, $\ldots=$const.), the larger N_N is*), the larger $\sigma_{0\,min}$ is.

b. — The influence of temperature and voltage

Realizing that n_i and Λ_{Di} are dependant on temperature but that $\sigma_0\neq f(T)$, we would better rewrite (2.26g) in the form:

(2.27a)
$$\sigma_{0\,min}(T)=\left(\frac{e^2N_N{}^3}{3\varepsilon\varepsilon_0}\right)^{1/2}\sqrt{\frac{1}{kT\sinh^{-1}\frac{N_N}{2n_i}}},$$

using (2.19) for Λ_{Di}, $\sigma_{0\,min}$ is clearly shown to have the *minimum* with respect to T; N_N being constant:

(2.27b)
$$\sigma_{0\,min\,min}=\frac{4n_i}{\Lambda_{Di}}\left[\frac{kT\left(\frac{N_N}{2n_i}\right)^2\sqrt{1+\left(\frac{N_N}{2n_i}\right)^2}}{9kT+3Wg}\right]^{1/2}$$

for $T=\Theta_0$ (and n_{i0}) which is obtained from the condition:

$$\sinh^{-1}\frac{N_N}{2n_i}=\frac{\frac{N_N}{2n_i}(3kT+W_g)}{2kT\left[1+\left(\frac{N_N}{2n_i}\right)^2\right]^{1/2}},\qquad n_i=\sqrt{B_cB_v}\,e^{-W_g/2kT}\equiv$$

(2.27c)
$$\equiv A(kT)^{3/2}e^{-\tilde{W}_g/2};\quad B_{c,v}=2(2\pi m_{gn,p}kTh^{-2})^{3/2}\sim(kT)^{3/2}.$$

*) $y(N_N)$ is a monotonously increasing function of N_N. The wrong conclusion might be obtained from the final right hand expression in (2.26g), i.e. that there is minimum of $\sqrt[3]{e}/\Lambda_{Di}$ in $y(N_N)$ for $N_N/n_i=\sqrt[3]{e}$ (while it is valid only for $N_N/n_i\gg 1$); therefore the relation (2.21) in [24] is not always correct, as $e^x\approx 2\sinh x$ do not hold always.

Consequently, the space charge limit (x_N) in the given transition may be *in the transition* ($x_N < x_2$) or *outside*:

$$\boxed{(2.27\,\mathrm{d})} \qquad \begin{aligned} x_N &\leqslant x_2 \quad \text{for} \quad \mathcal{S}_{0\,min}(T_1) \geqslant \mathcal{S}_0, \\ x_N &\geqslant x_2 \quad \text{for} \quad \mathcal{S}_{0\,min}(T_2) \leqslant \mathcal{S}_0, \end{aligned}$$

at various temperatures. The fact that $\mathcal{S}_{0\,min}$ *increases* for $T > \Theta_0$ if T is increased, and it *decreases* for $T < \Theta$, is even more important. It actually means that x_N in the former case decreases and it increases in the latter (it may even ,,get out" of the transition).

In other words, x_N either diminishes (if $T > \Theta_N$) or it enlarges (if $T < \Theta_N$) when T increases — actually it has its *maximum* for $T = \Theta_N$ (Θ_N is different from Θ_0) as it will be shown more generally in Sec. 2.4. both for the linear and for the abrupt transition.

When the *applied voltage* U is sufficiently "small" so that the problem can be solved even further electrostatically, introducing the notion of quasi-Fermi levels as in (2.19), when it is very convenient to introduce the notions of *effective* Λ_D's and $\mathcal{N}$'s:

$$\text{(2.28 a)} \qquad \Lambda_{Di\,U} \equiv \Lambda_{DU} = \Lambda_{Di}\, e^{-e\tilde{U}/4}, \qquad \mathcal{N}_U(\boldsymbol{r}) = \mathcal{N}(\boldsymbol{r})\, e^{-\frac{e\tilde{U}}{2}}$$

and the notion of *effective* slope and ρ's:

$$\text{(2.28 b)} \qquad \mathcal{S}_{0\,U} \equiv \mathcal{S} = \mathcal{S}_0\, e^{-\frac{e\tilde{U}}{2}} \quad \text{and} \quad \rho_U = \rho\, e^{-e\tilde{U}/2},$$

which all *depend* on U. That *physically* means that all processes happen so as if $U = 0$, but with the *effective intrinsic* concentration:

$$\text{(2.28 c)} \qquad n_{iU} = n_i\, e^{e\tilde{U}/2} \equiv n_i\, \beta, \quad (\beta = e^{e\tilde{U}/2}),$$

or rather — as if the gap W_g had been changed for eU and *all* other values had remained the same [24].

If we compare (2.28 c) with the results obtained under the assumption that the condition $\mu_{fp} = \frac{eU}{2} = \mu_{fn}$ has to be satisfied (for instance as in [9] — page 243), we see that this condition is not necessary.

Consequently, whatever is derived in Section (*a*) is valid also for $U \neq 0$ (for smaller polarization). Thus the condition (2.26 b) changes into:

$$\boxed{(2.28\,\mathrm{d})} \qquad \left(\frac{\Lambda_0}{\Lambda_{Di}}\right)^2 \beta^2 \geqslant \frac{1}{8} \quad \text{or} \quad \Lambda_0\, \beta \geqslant \Lambda_{Di},$$

from which the direct polarizations are seen to ,,help" the quasi-neutrality appearance. Above all, voltage application can *completely change* the condition for depletion (or for the quasi-neutrality). Direct polarization influences (but not those of reverse polarization) are also examined in another way in]26], and the same conclusions are attained just like with us.

$\mathcal{S}_{0\,min}(T, U)$ analysis leads to the *conclusion* (which is deducted in the original version [21]) that $\mathcal{S}_{0\,min}$ again has the *minimum* for $T = \Theta_{0U}$ which *increases*

when U_{dir} *increases* and it *decreases* when U_{rev} *increases*, soon tending towards the asymptotic value

$$k\Theta_{0a}=\frac{1}{e}\left(\frac{N_N}{A}\right)^{3/2},\quad A=2\left(\frac{2\pi}{h^2}\right)^{3/2}(m_{gn}\,m_{gp})^{3/4}. \tag{2.28e}$$

The dependence of SC- width upon T and U will be analysed in Sec. 2.4. where the estimation of values Θ_{0a} for $\mathcal{S}_{0U}$ will be given.

c. — Recombination process influence; transition capacitance

1° — It is comprehensible that the recombination processes in the transition have not any influence when they are in the equilibrium state (without polarization), but as soon as nonequilibrium concentrations (the polarizations $U\neq 0$) appear, these processes may ,,spoil'' *rectifying* action and they can affect the distribution of φ and ρ very much. In principle, the complete system of equations (2.2), (2.3) and (2.4) or (2.19) should be solved, and for the steady state the condition (2.5) is satistied.

Shockey accepted the simplest model of *direct* recombination (instead of r_e according to eqn. 2.7, one has $r_e=\gamma_r np-g_i$) and he got:

$$J\approx en_i^2\,\gamma_r\int_{l_P}^{l_N}(e^{\tilde{E}_{fn}-\tilde{E}_{fp}}-1)\,dx\equiv en_i^2\,\gamma_r\cdot\mathcal{J}_f \tag{2.29a}$$

for the total current, where l_N and l_P are points of contacts sufficiently spaced.

$\mathcal{J}_f$ serves as the measure for the total number of carriers which must be recombined in the transition in order to make J at l_P and l_N to be composed almost entirely of majority carriers. It means that for the constant J and T, the larger the recombination (γ_r), the smaller the difference ($E_{fu}-E_{fp}$) must be.

— When the recombination is *extremely intensive*, quasi-Fermi-levels join and

$$U=\frac{E_{fp}(l_N)-E_{fn}(l_P)}{e}=J\int_{l_N}^{l_P}\frac{dx}{\sigma}\equiv JS\cdot R_0\equiv IR_0, \tag{2.29b}$$

is obtained, where $\sigma=\sigma_n+\sigma_p$ is purely ,,ohmic'' (drift) conductivity, so that $U=IR_0$ is the ohmic voltage drop: *rectification* practically does not exist*). $\gamma_r\to\infty$

— $E_{fn}(x)$ and $E_{fp}(x)$ differ considerably if the *recombination is moderate*, and by introducing the notion of *transition voltage* $U_i\cdot e=E_{fn}(0)-E_{fp}(0)$ according to [24] (we consider that it would be more general to define U_i for $x=x_i$ and not for $x=0$) and supposing that $E_{fn}(x)-E_{fp}(x)=eU_i=const$ for

*) As p and n depend upon x, and hence upon U as well, at least on the width Λ_ρ of transition, then however R_0 will depend upon U at least slightly and the linearity U versus I will not be complete: there is a certain rectification, which is not always stressed. Consequently, the *term* ,,ohmic'' in the sense of drift and the term "rectifying" should be *differenciated.*

the whole integration region ($L_{pN} \geqslant x \geqslant -L_{np}$) of the diffusion lengths beyond which $E_{fn} - E_{fp} = 0$, one has:

$$J = en_i^2 \gamma_r (L_{pN} + L_{nP}) (e^{e\tilde{U}_i} - 1) \equiv J_S (e^{e\tilde{U}_i} - 1), \tag{2.29 c}$$

where J_S is the total number of recombined pairs in the layer $(L_{nP} + L_{pN})$ or the saturation current. The remainder $(U - U_i)$ of the voltage drop is ohmic:

$$U - U_i = \delta U_{Rp} + \delta U_{Rn} \equiv \int_{l_P}^{x_i} I_p \frac{dx}{\sigma_P} + \int_{x_i}^{l_N} I_n \frac{dx}{\sigma_n} \equiv IR_1 \tag{2.29 d}$$

so that finally

$$U = U_i + \delta U_{Rp} + \delta U_{Rn} = R_1 I + \frac{kT}{e} \ln\left(1 + \frac{I}{I_s}\right) \tag{2.29 e}$$

is obtained, and evidently $\frac{R_0}{2} < R_1 < R_0$. Consequently, the recombination involves the series resistance R_1, and it influences the transition state just because it *reduces* U to U_i.

2° — The basic characteristic of every transition is its *capacitance* C_{tr}, or more generally expressed — its *susceptance*. We should mention first that there are *two conceptions* of capacitance, the well known „electrostatic":

$$C_{tr} = C_\rho \equiv -\frac{dQ}{dU} \equiv \frac{\partial}{\partial U} \int_{l_P}^{x_{Kmax}} S \rho \, dx, \quad \left(\int_{l_P}^{l_N} \rho \, dx = 0\right) \tag{2.30 a}$$

and the almost unnoticed *dynamic conception* according to Adirovič ([23]—1959):

$$C_{tr} = C_\omega \equiv eS \int_{l_P}^{l_N} \frac{\partial p}{\partial U} dx = eS \int_{l_P}^{l_N} \frac{\partial n}{\partial U} dx, \tag{2.30 b}$$

which seems to be more adequate*).

The transition capacitance for the „neutral" case ($\mathcal{S}_0 < \mathcal{S}_i \equiv n_i/\Lambda_{Di}$) is examined in detail in [21]:

$$\boxed{(2.30\,c)} \quad \frac{C_{neut}}{S} = e \frac{d}{dU} \int_{l_P}^{l_N} p \, dx = \cdots = \frac{C_\omega}{S} - e \frac{dl_P}{dU} p(l_P) \cdot \left[1 + \frac{\left|\frac{dl_N}{dU}\right| \cdot p(l_N)}{\left|\frac{dl_P}{dU}\right| \cdot p(l_P)}\right]$$

$\left(\text{from which it is concluded that } C_\omega \geqslant C_{neut}, \text{ because always } \frac{dl_P}{dU} \geqslant 0 \text{ holds}\right)$, as well as the recombination effect on C_{tr} in general. There, by dividing the the current into the time independent (J_{on}) and the time dependent part ($J_{\omega n}$),

*) This question is discussed in more detail in [21]. See also [45], where there is an attempt to evaluate the difference between C_ω, C_ρ and $^\circ C_\rho$ — capacitance defined in (2.30a) but in the approximation of total depletion. We remark that $^\circ C_\rho$ but not C_ρ, agrees with the experiment.

for dc-part we have got the losses R_0 and R_1 like in (29 b and d) and the excess concentration of time dependent part

$$\underline{\delta n_\omega} = \delta n_\omega\,(0)\,e^{j\omega t}\,\frac{1-\exp\dfrac{x-l_N}{\underline{L_K^*}}}{1-\exp\left(-\dfrac{l_N}{\underline{L_K^*}}\right)},$$

(2.30 d)

$$\frac{1}{\underline{L_K^*}} = \frac{\sqrt{1+j\omega\tau}}{L_K^*} \equiv \frac{1}{\mathscr{L}_K}(1+j\xi),\ (j=\sqrt{-1}),$$

$$L_K^* = \frac{2D}{(\mu^2K^2+4D/\tau)^{1/2}+\mu K}.$$

We see that $\underline{\delta n_\omega} \to \mathrm{Re}\,\underline{\delta n\omega}$ *only* when $\omega\tau \to O$, which means that the current will be purely reactive ($\underline{j\omega C_\omega U}$) only for frequencies sufficiently low and for the recombination sufficiently weak (short lifetime τ). Otherwise, by writing $\underline{\delta n_\omega} = a + jb$, the *susceptance* has the form derived in [21]:

(2.30 e)

$$\frac{\underline{U(G+j\omega C_\omega)}}{S} \equiv \left[e\int_{l_P}^{l_N}\frac{a}{\tau}\,dx - e\,\omega\,U\int_{l_P}^{l_N}\frac{\partial b}{\partial U}\,dx\right] + j\left[e\int_{l_P}^{l_N}\frac{b}{\tau}\,dx + e\,\omega\,U\int_{l_P}^{l_N}\frac{\partial a}{\partial U}\,dx\right].$$

It means that the recombination *influences* both G and $C_\omega = C_{tr}$, with the effect that both the values *depend* upon frequency, recombination mechanism, polarization etc.

Let us remark that the expressions (2.30 d and e) have a truly general character and that we have not had any opportunity to see them at any other place.

2.4. — Temperature dependences of Λ_ρ and φ

Widths of SC-region (Λ_ρ or x_N and x_P) are shown here analytically to have the *maximum* for some temperature (Θ). Temperature variation of the potential distribution $\varphi(x)$ is analysed.

Experimental checking will be carried out in Chapter 4.

a. — Starting relations for temperature dependence of the SC-region width

In the general case ($U \neq 0 \neq U'$, see also the condition 2.31 e) *defining* $x_{N,P}$ with $\rho \approx 0$ in (2.19):

(2.31 a)
$$\sinh\tilde{\psi}_{N,P} \approx \frac{\mathscr{N}(x_{N,P})}{2\beta},\quad \beta = e^{\frac{e\tilde{U}'}{2}} = const,$$

under the assumption that the quasi-neutrality condition is valid here as well as that $\psi_{N,P}$ and $n_i \subset \mathscr{N}(x_{N,P})$ explicitly depend *only* upon $x_{N,P}$, and (βn_i) *only*

on T, considering that W_g, ε and the like do not depend on T, and differentiating both sides of (2.31a) with respect to (kT) (here the sign „prim")

$$\boxed{(2.31\,b)} \qquad x'_{N,P} \equiv \frac{dx_{N,P}}{dkT} = \ldots = \frac{2\left(\frac{\beta n_i}{kT}\right)^2 \cosh\tilde{\psi}_{N,P}\cdot\psi_{N,P} - \mathcal{N}(x_{N,P})\cdot n_i\cdot(\beta n_i)'}{2\frac{(\beta n_i)^2}{kT}\cosh\tilde{\psi}_{N,P}\cdot\frac{d\psi_{N,P}}{dx_{N,P}} - \mathcal{S}_0(x_{N,P})\cdot\beta n_i},$$

$$(2.31\,c) \qquad (\beta n_i)' = \frac{\beta\tilde{n}_i}{2}(3+\tilde{W}_g - e\tilde{U}'); \quad n_i\beta = A(kT)^{3/2}\,e^{-\frac{W_g - eU'}{2kT}}$$

is easily obtained, with all the symbols as before.

The denominater in (31b) can be proved never to be zero (for the abrupt transition it is simply $\mathcal{S}_0 = 0$) so that $x_{N,P}$ have the *extremum* for $x'_{N,P} = 0$:

$$\boxed{(2.31\,d)} \quad 2\left(\frac{\beta n_i}{kT}\right)^2 \cosh\tilde{\psi}_{N,P}\cdot\psi_{N,P} = \mathcal{N}_{N,P}\cdot n_i\cdot(\beta n_i)' = 2\beta n_i\cdot\sinh\tilde{\psi}_{N,P}\cdot(\beta n_i)',$$

$$\therefore \tilde{\psi}_{N,P} = \tanh\tilde{\psi}_{N,P}\cdot\frac{(\beta n_i)'}{\beta\tilde{n}_i} = \tanh\tilde{\psi}_{N,P}\cdot\left(\frac{3}{2}+\frac{\tilde{W}_g - e\tilde{U}}{2}\right).$$

Using (31a and c) and taking into account (31d), the $x'_{N,P}$ from (31b) is seen to be:

— negative $\left(\frac{n_i\mathcal{N}}{\beta n_i}, \tilde{W}_g, \tilde{U}'\right.$ and $\psi_{N,P}\to 0$, while $\left.\tilde{\psi}_{N,P}\cdot\coth\tilde{\psi}_{N,P}\to 1\right)$ for the *high* T's $\left(\text{formally } T\to\infty, \text{ so that } 1-\frac{3}{2} = -\frac{1}{2} < 0\right)$

— but it is *positive* $\left(\frac{\mathcal{N}}{2\beta} \gg 1, \sinh\tilde{\psi}_{N,P} \approx e^{\tilde{\psi}_{N,P}}/2\right)$ for the *low* T's $\left(\text{formally } T\to 0, \text{ and as } \tilde{\psi}_{N,P} \approx \ln\frac{\mathcal{N} n_i}{A} + \frac{W_g - eU'}{2kT} + \frac{3}{2}\ln\frac{1}{kT}\right.$ and $\tanh\tilde{\psi}_{N,P}\to 1$ then $\left.c^{te} - \frac{3}{2} + \ln\frac{1}{kT}\to +\infty\right)$ it means that $x_{N,P}$ have the *maximum* values ($x_{N\,max}$ and $x_{P\,max}$) for the *temperatures* $T = \Theta_{N,PU'}$. It may be defined from the transcedental equations (31d and a). If the conditions

$$(2.31\,e) \qquad \left|\frac{\mathcal{N}(x_{N,P\,max})}{2\beta}\right| \equiv \left|\frac{n_i\mathcal{N}}{2\beta n_i}\right| \equiv \frac{|n_i\mathcal{N}|}{2A(k\Theta)^{3/2}}\cdot e^{\frac{W_g - eU'}{2k\Theta}} = |\sinh\psi_{N,P\,max}| \gg 1,$$

$$\therefore \frac{W_g - eU'}{2kT} > \text{or} \gg 1; \quad |\sinh\tilde{\psi}| \approx \frac{e^{|\tilde{\psi}|}}{2}; \quad |\tanh\tilde{\psi}| \approx 1 - \frac{1}{2\sinh^2\tilde{\psi}}$$

are supposed to be fulfilled, then from (31d and a)

$$|\sinh\psi| = \frac{|n_i\mathcal{N}|}{2\beta n_i} = \frac{|n_i\mathcal{N}|}{2A(k\Theta)^{3/2}}\cdot e^{\frac{W_g - eU'}{2k\Theta}} = \frac{1}{2}e^{|\tanh\tilde{\psi}|\left(\frac{3}{2}+\frac{W_g - eU'}{2k\Theta}\right)},$$

$$\boxed{(2.31\,f)} \qquad \therefore \frac{|\psi(x_{N,P\,max})|}{k\Theta_{N,PU}} \approx \frac{3}{2} + \frac{W_g - eU'}{2k\Theta_{N,PU}}, \quad (\text{for } |\tanh\tilde{\psi}| \approx 1);$$

$$(k\Theta_{N,PU})^{3/2} = \frac{|n_i\cdot\mathcal{N}(x_{N,P\,max})|}{Ae^{3/2}}\left\{1 + \left[\left|\frac{\beta n_i}{\mathcal{N} n_i}\right|^2_{T=\Theta_{N,PU}}\cdot\left(3 + \frac{W_g - eU'}{k\Theta_{N,PU}}\right)\right]\right\},$$

are obtained; the member [...] is much smaller than 1.

Here the expressions (31 b, d and f) are more general than the corresponding ones in [21]; there they are given for particular cases of abrupt and graded transitions.

b. — The analysis of x_{NU} (or Λ_ϱ) and $\mathcal{S}_{0U}$ temperature dependences

The following may be concluded on the basis of the above mentioned:

— according to (31 f)-above, supposing that ψ depends only upon $x_{N,P}$ (explicitly) with the increase of the reverse voltage (with the absolute value), $x_{N,Pmax}$ *ceases* to depend explicitly upon $\Theta_{N,PU}$, and that happens as soon as $3 \ll \frac{W_g - eU'}{k\,\Theta_{N,PU}}$, what is *practically always* fulfilled (except for $W_g \approx eU'$ — the high injection level), because $\Theta_{N,PU}$'s are very low temperatures, as we shall see later on;

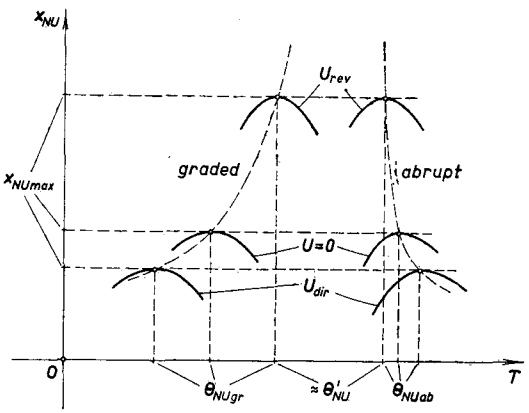

Fig. 2.4. — Temperature dependences of the space charge width for the graded (left) and abrupt transition (right), with U as parameter.

— therefore, the *characteristic temperature* $\Theta_{N,PU}$ will either *increase*, if $\mathcal{N}(x_{N,Pmax})$ grows more rapidly than $\{1+[\cdots]\}$ decreases — in the cases of graded transitions, or it will *slightly decrease* — in the case of abrupt transitions (where $\mathcal{N}=\text{const}$) if $|U_{inv}|$ is *increased*, according to eqn. (31 f) — down since $[\cdots]$ *always* decreases.

— Θ_{NU} and Θ_{PU} are the same only for the symmetric transitions;

— $\Theta_{N,PU}$ *depend very strongly* upon $|n_i \mathcal{N}(x_{N,Pmax})|$ i.e. upon the doping level of the region, rising with the increase of the latter.

The shapes of $x_{N,PU}(T)$ with U as the parameter are shown in Fig. 2.4., so that:

$$(2.31\,g) \qquad (k\,\Theta'_{N,PU})^{3/2} = \frac{|n_i \cdot \mathcal{N}(x_{N,Pmax})|}{A e^{3/2}}, \quad \left[A = 2\left(\frac{2\pi}{h}\right)^{3/2}(m_{gn}\,m_{gp})^{3/4}\right]$$

is the ,,asymptotic'' value.

As an *example*, let us calculate Θ_{0U} and Θ_{NU} for N-region of P-N junction of Ge and Si, taking $N = 2 \cdot 10^{16}\ \text{cm}^{-3}$ what would correspond to the Ge of the conductance of about 10 S/cm or about 3 S/cm for Si, at the ambient temperature (typical value for the base conductivity both for the ,,drawn'' and for the alloyed germanium transistor transitions); let us take that $N = N_N$ — the final concentration for the graded transition, when we calculate $\mathcal{S}_{0min}$ and $\mathcal{S}_{0minU}$ i.e. — Θ's and let us take that its ,,real'' physical slope is $\mathcal{S}_0 = 10^{16}\ \text{cm}^{-3}/\text{cm}$ (the value which is given for Si transition [24]) and $N = C^{te}$ — the donor contents for the abrupt transition, which is supposed to be symmetrical; the *results* are shown in T 2.2. when the values for W_g and n_i (300 °K) are taken from [13] and they are 0.72 and 1.1eV, or 2.4×10^{13} and $1.4 \times 10^{10}\ \text{cm}^{-3}$ for Ge and Si, respectively. In order to see the *application limits* of the cor-

responding appoximative expressions, the whole account is *repeated* for $N=2\times10^{19}$ cm^{-3}, what approximatively represents the boundary of the non-degenerated carrier states (for the ambient temperature).

The following may be *concluded* from T 2.2:

— The characteristic temperatures Θ_{0a} and Θ_{0U} for $\mathcal{S}_0$'s are low even for the case when N is almost equal to the *effective density of states* ($B_{c,v}$) at the ambient temperature (then the semiconductor is already degenerated), and apart from that, they are practically *independent* of U ($\Theta_{0U}\approx\Theta_U\approx\Theta_{0a}$) even up to $99\,W_g/100e$ of the direct polarization; for the adopted $\mathcal{S}_0$, there are no real reverse voltages which will widen the depleted space in the region $N=N_N=const.$, when $N_N=2\times10^{16}-10^{19}$ cm^{-3}.

— Θ_{NU} are *still lower* for the *graded transition*, so that the depletion region boundary (x_{NU}) decreases with the increase of T for all real temperature and U's (see Fig. 2.4); here

$$x_{NU\,max}\approx[3\,\varepsilon\varepsilon_0\,(W_g-eU)/2\,e^2\,\mathcal{S}_0]^{1/3}\sim\sqrt[3]{W_g-eU}. \tag{2.31 h}$$

— x_{NU}'s are smaller in the *abrupt* (symmetrical) transition, when all other conditions are the same, what is otherwise quite clear but we should keep in mind that Θ_{NU} *increases* when the direct polarization increases (but very slightly) and it *decreases* for the inverse polarization (also slightly) so that $\Theta_{NU}\approx\Theta_{0a}=\Theta'_{NU}$ as in Fig. 2.4.; otherwise the following is obtained from the relation (2.31f) for such a transition:

$$x_{NU\,max}=\left(\frac{\varepsilon\varepsilon_0}{e^2N}\,W_{gU}\right)^{1/2},\quad W_{gU}=W_g-eU, \tag{2.31 i}$$

$$(k\,\Theta_{abU})^{3/2}=\frac{\sqrt{PN}}{A\,e^{3/2}}\left[1+\frac{n_i^2\,e^{e\bar{U}}}{2PN}\left(\frac{P}{N}+\frac{N}{P}\right)\left(3+\frac{W_{gU}}{k\,\Theta_{abU}}\right)\right], \tag{2.31 j}$$

where the last expression is valid also for the asymmetrical abrupt transition ($P\neq N$).

The starting point for the whole this analysis is the assumption that all the states are quasi-equilibrium states, so that the above conclusions are quite correct within that frame, because we „must not" apply too high reverse voltages, nor „permit" large currents flow, i.e. $eU\ll W_g$ for the direct direction, when the conditions (2.13e) — below are *certainly fulfilled.* Thus the value of the voltage larger than W_g/e in the last row of T2.2 must be understood to be only conditional — i.e. only as the reliable indication that such Θ cannot be obtained.

c. — Potential temperature dependence for the abrupt transition

The dependence $\varphi(T)$ for the more graded symmetrical transition can be analysed comparatively easily. The accurate analysis of the asymmetric abrupt transitions is rather hampered by the ignorance of the exact expression for x_i (for the symbols see also Fig. 2.3b) and therefore its dependence upon T is not known.

Characteristic temperatures (Θ) *and others for the concrete Ge and Si transitions* T 2.2

Quantity [unit]	for $N = 2 \times 10^{16}$ cm^{-3} Ge	for $N = 2 \times 10^{16}$ cm^{-3} Si	for $N = 2 \times 10^{19}$ cm^{-3} Ge	for $N = 2 \times 10^{19}$ cm^{-3} Si	calculated from
$A \times 10^{-20}$ see also (2.31 g) $\left[\frac{\text{cm}^{-3}}{(\text{eV})^{3/2}}\right]$	75	60	&)	&)	(2.31c)
ε — relative diel. cont.	16	&)	&)	12	—
Θ_{0a} — asympt. temp. for $\mathcal{S}_{0\,min\,min}$ [°K]	0.77	0.96	77	96	(2.28e)
$\frac{n_i^2 e^{e\tilde{U}}}{N_N^2}\left(3+\frac{Wg-eU}{k\Theta_{0a}}\right)\approx$ for $U=0$ and	$\ll 10^{-46}$	$\ll 10^{-59}$	10^{-46}	10^{-60}	(2.31f)
$eU = 0{,}99\,Wg$	$\ll 10^{-3}$	$\ll 10^{-3}$	0.062	0.056	
Necessary slope [10^{20} cm^{-4}] $\mathcal{S}_{0\,min}$ for $T = 300$°K	12.9	10.4	294	275	(2.27a)
and $\mathcal{S}_{0\,min\,min}$ for Θ_{0g}	9.2	8.5	292	272	(2.27b)
U_{rev} necessary to equate [GV] $\mathcal{S}_{0\,min\,U}$ with $\mathcal{S}_0$ for 300°K	2.64	3.96	3040	4130	(2.28c)
for Θ_{0a}	6.1	8.0	6150	8100	(2.28c)
Graded transition, $U=0$: $\Theta'_{NU} \times 10^4$—for $x_{N\,max}$ [°K]	14	&)	&)	72	(2.31g)
x_{NU} for 300°K depletion region width [μm]	1.4	&)	&)	7.1	
x_{NUmax} for Θ'_N	9.8	&)	&)	10	(2.31h)
Abrupt transition, $U=0$: $\Theta'_N \approx \Theta_{0a}$ for $x_{N\,max}$ [°K]	0.77	0.96	77	96	(2.31g)
X_{NU} 300°K — boundary [nm]	124	156	1.78	2.00	(2.31i)
$X_{NU\,max}$ Θ_N	178	190	5.60	6.00	
eU to be $k\Theta_{NU} = W_d$*) with graded transition	-10^{27}	-10^{28}	&)	&)	(2.31h)
[meV] abrupt transition: Wd Ge; Si ≈ 0,01; 0,04 eV $Wg+$	+3.3	+11.4	+0.52	+4.7	**)

&) The value does not depend upon N

$G = 10^9$ (giga).

1 μm = 10^{-6} m = 10^{-4} cm = 10^{-3} mm;
1 nm = 10^{-9} m = 10^{-3} μm.

*) W_d — Ionization energy of donor impurity, chosen so that it corresponds to that of antimony as donor [13]; those are temperatures $\approx W_d/2k = 58$°K for Ge and ≈ 230°K for Si, above which this impurity is almost completely ionized; the sign „—" corresponds to the reverse polarization and „+" to the direct polarization.

**) These voltages cannot be calculated from the approximate expressions of the type (2.31f) or (2.31j) (with $P=N$), because the „correction" for $eU > Wg$ is $[\ldots] \gg 1$, so that the correct expression (1.27a) from [21] with $x'_{NU}=0$ (and $P=N$) is used and the graphical solution is performed: drawing $\left[\eta_U^{-1}\left(1+\eta_U^2\right)^{1/2}\right]\cdot\sinh^{-1}\eta_U$ and

$$\left[\frac{3}{2}+\frac{W_{gU}}{2kT}\right]=\left[\frac{3}{2}+\ln\frac{2A(kT)^{3/2}}{N}+\ln\eta_U\right] \text{ versus } \eta_U \equiv \frac{N}{2n_{iU}},\ (kT = k\Theta_{NU} = Wd).$$

The dependence of $\varphi_{N,P}$ upon T can be determined exactly (according to the eqn. 2.31a) and it is shown that both dependences have the *maximum* (φ_N and $|\varphi_P|$) with respect to T for:

$$(k\Theta_{U\varphi_{N;P}})^{3/2}=\frac{N;P}{A\mathrm{e}^{3/2}}\left[1+\frac{n_{iU}^2}{N^2;P^2}\left(3+\frac{W_{gU}}{k\Theta_{U\varphi_{N;P}}}\right)\right]. \tag{2.32a}$$

Taking the risk of not getting quite adequate results, let us accept Sh--approximation. Then, as it is well known, one has

$$^{\circ}\tilde{\varphi}^{\circ}=\frac{\tilde{\varphi}_N N+\tilde{\varphi}_P P}{N+P}+\left\{\frac{2n_i}{P+N}(\cosh\tilde{\varphi}_P-\cosh\tilde{\varphi}_N)\right\} \tag{2.32b}$$

(The above left hand side index means zero i.e. the Sh-approximation; if the expression in brackets is added, the exact value $\tilde{\varphi}^{\circ}$ is obtained), and $^{\circ}x_i$ will be then given by the relation in TB.e. On the basis of these two relations, it is easy to prove that both $^{\circ}\varphi^{\circ}(T)$ and $^{\circ}x_i(T)$ have the *maxima* for:

$$(k\Theta_{U\varphi^{\circ}})^{3/2}=\frac{1}{A\mathrm{e}^{3/2}}\cdot\frac{P^{P/P-N}}{N^{N/P-N}}\left[1-\frac{n_{iU}^2}{PN}\left(3+\frac{W_{gU}}{k\Theta_{U\varphi^{\circ}}}\right)\right];$$

$$\frac{\Theta'_{x_i U}}{\Theta'_{abU}}\approx\left(\frac{P}{N}\right)^{0{,}806}\quad\text{for } P\gg N. \tag{2.32c and d}$$

It is necessary to notice that Θ for φ° *decreases* (slightly) with an increase of U, like with a graded transition (see also Fig. 2.4 — on the left) and that the *asymptotic value* Θ' for the inversion point (region boundary enriched with holes) may be *considerably larger* than that for the transition width.

d. — The conclusion for temperature dependences

According to the above mentioned, the conclusion would be:

— that the SC-region widths have *maxima* as indicated in Fig. 2.4, both for the graded and for the abrupt transition,

— that the dependences of the slope $\mathcal{S}_{0\min U}(T)$ have the *minimum* (Fig. 2.5a),

— that the potential $\varphi(x)$ moves as in Fig. 2.5b, where the dotted line denotes the potential distribution for the *pseudo-abrupt* transition treated in detail in [21], where, apart from other things, it is proved that always $\varphi_{ps}<\varphi_{ab}$ and $x_i<x_{ips}$, if all other conditions are the same.

Interesting temperature effects, not noticed by the others, are theoretically proved in this way. It is necessary to emphasize that the same shape (with the maximum) has also *contact potential difference* ($eU_D=\varphi_N-\varphi_P$) which actually defines the barrier in the way physically most adequate and most objective, since the notions of Λ_ρ or x_N and x_P are conditional to a very great rate: where the SC-region ceases is not known precisely.

All these results are *obtained under the condition* that the impurity centres are *completely ionized*, and that Θ's are low, which is evidently contradictory.

Nevertheless, it is shown in one of the latest papers [44], both *experimentally and theoretically*, that the temperature maximum effect always exists regardless the degree of impurity ionization.

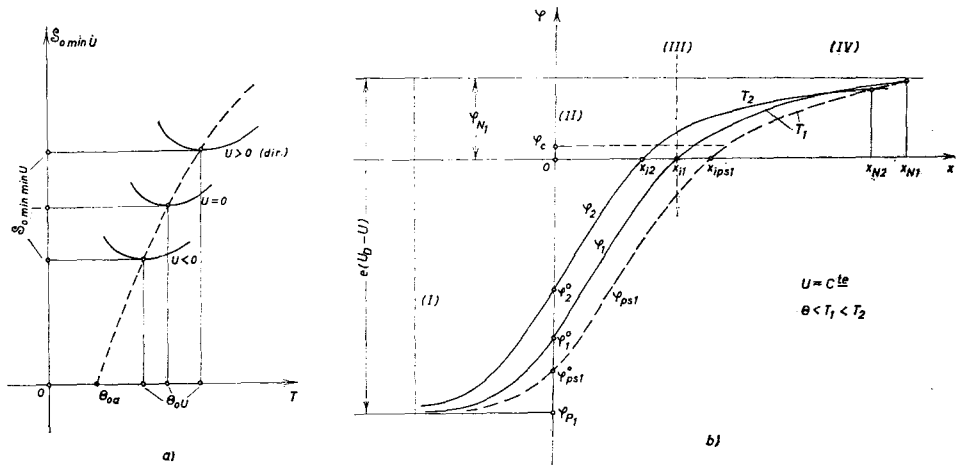

Fig. 2.5 — Temperature dependences for the least transition slope required (*a*) and $\varphi(x, T)$ — the abrupt transition potential (*b*); subscripts "1" and "2" refer to T_1 and T_2, "*ps*" — pseudo-abrupt transition.

Chapter 3. — LIMITING REGIMES IN P-N STRUCTURES

Limiting regimes are understood to be either direct polarization currents caused by the polarizations of the order U_D or reverse polarizations which lead to the (avalanche) breakdown with the currents that overtake the saturation currents (I_S) for many orders of magnitude.

Considering that all the impurity centres are ionized, let us remind of the conditions that exist in the transition regions:

I. — Carrier depletion — impurities have got bare,

II. — The field is nonhomogeneous and rather strong, even when the polarization is direct (10^3—10^4 V/cm), so that the mobility (μ) *depends* upon K (at least near K_{max} at $x=0$).

III. — With the increase of K, the (avalanche) breakdown may occur (see for instance [28]). Because of the finite value of the width Λ_ϱ, this breakdown need not be the same as that in a homogeneous body.

3.1. Direct polarization limiting regimes

According to the generally known conception, let us assume that the direct current is composed of the *diffusion* current (J_d) on the boudaries x_N^* and x_P^* (see also Fig. 2.1.) and the *recombination-generation* current (J_{g-r}) in the transition:

$$J_{dir} = J_d + J_{g-r}; \qquad J_d = J_{dp} + J_{dn}. \tag{3.1.}$$

Let us consider first J_{g-r} and then J_d — roughly and generally.

a. — Recombination-generation currents from the transition

1° — Taking into consideration only *analytical solutions* which, in our opinion, are given in the best way in [32] and especially in [22], for the recombination — generation part J_{g-r} of the total current J_{dir}, the following is obtained:

$$(3.2a)\quad J_{g-r}=e\left[\frac{1+\frac{x_N}{L_{nP}}\cdot I_{1n}(U')}{1+\frac{x_N}{L_{nP}}\cdot I_{2n}(U')}+\frac{1+\frac{x_N}{L_{pN}}\cdot I_{1p}(U')}{1+\frac{x_N}{L_{pN}}\cdot I_{2p}(U')}\right]\cdot I_3^*(U')\cdot\frac{n_i x_N}{\tau}\cdot 2\sinh\left(\frac{eU'}{2kT}\right),$$

so that according to [22] the integrals I_i $(i=1, 2, 3)$ are defined *only* for the *linear graded symmetrical* transition and that as (see also eqn. 2.17c)

$$(3.2b)\quad I_{1n,p}(U')=2e^{-\alpha^*}\int_0^1\gamma_{\mu n,p}\cdot t\cdot\sinh\tilde{\varphi}\,dt,\quad \tilde{\varphi}=\frac{\alpha^*}{2}(3t-t^3);$$

$$I_{2n,p}(U')=2\,e^{-\alpha^*}\int_0^1\gamma_{\mu n,p}\cdot\cosh\tilde{\varphi}\,dt,\quad \alpha^*\equiv\frac{e\tilde{U}^+}{2}=\alpha+\ln\frac{n_N^\circ}{N};$$

$$(3.2c)\quad I_3^*(U')=\int_0^1\frac{dt}{2\cosh\tilde{\varphi}},\quad\left(t=\frac{x}{x_N^*}\right)$$

but with $U'=U$, $\alpha^*=\alpha\equiv e\frac{U_D-U}{2kT}$, $\gamma_{\mu n,p}=1$ and $\cosh\tilde{\varphi}\approx e^{\tilde{\varphi}}/2$. Consequently, even in [22] the attention was not paid, that:

(I) — the total voltage (U) is not the same as the voltage at the transition $(U'$ — see also Fig. 2.1.),

(II) the „correction" according to (2.16) and (2.17c) may be considerable especially at higher levels of injection, and that it *is not* $\alpha^*=\alpha$;

(III) that $\gamma_{\mu n,p}\equiv\frac{\mu_{0n,p}}{\mu_{n,p}(K)}$ — the ratio between the starting (for $K=0$) and actual mobility may differ considerably from the unity, and

(IV) that the question of approximation $\cosh\tilde{\varphi}\approx e^{\tilde{\varphi}}/2$ is discussable — especially for the abrupt asymmetric transitions (which are not treated in [22]) and for the direct polarization (when $\tilde{\varphi}_{max}=\tilde{\varphi}_N$ or $|\tilde{\varphi}^\circ|$ is not much larger than 1).

The first remark is not essential, but the last two must be examined in the light of the second (II).

2° — *The integrals* I_1 *and* I_2. In [22] the starting point is the relation of the type (2.2) and (2.3) with $\frac{\partial n}{\partial t}=0$ and $g'_{n,p}=0$, as well as $r_{n,pe}$ according to (2.7 with $W_R\approx W_g/2$, what is quite correct, but it is considered that $\mu\neq f(K)$ i.e. that it does not depend upon x, so that from

$$(3.3a)\quad \frac{dJ_n}{dx}=er_e\;\therefore\;J_n=e\int r_e\,d_x+C_n=\mu_n(x)\cdot kT\left(\frac{dn}{dx}+n\frac{e}{kT}K\right),$$

the differential equation for n is obtained:

$$\frac{dn}{dx}+e\tilde{K}(x)\cdot n=\frac{e}{\mu_n(x)}\int \tilde{r}_e\,dx+\frac{C_n}{kT\cdot\mu_n(x)} \tag{3.3b}$$

and the solutions *differ* from those in [22] as the function $\gamma_{\mu n,p}(x)$ is also included in (3.2b). Apart from that, α^* should be placed instead of α, as it was taken in [22].

It can be easily shown [21] that the correction due to γ_μ may *increase* $I_{1,2}$'s mostly twice or four times, so that if $x_N \ll L$, then the values $I_{1,2}$ do *not influence* the exactness of the expression for J_{g-r}, the value $[\cdots]$ in (3.2a) is about 2 for „good" transitions.

3° — The integral I_3^* for the *symmetrical* graded linear transition defined with (3.2c) and with $\tilde{\varphi}$ from (3.2b) cannot be calculated in a simple way, and the following two ways can be applied:

— either $\operatorname{sech}\tilde{\varphi}$ is expanded into the series at $\tilde{\varphi}=a$:

$$\operatorname{sech}\tilde{\varphi}\equiv\frac{1}{\cosh\tilde{\varphi}}=\frac{1}{\cosh a}-\frac{\sinh a}{\cosh^2 a}(\tilde{\varphi}-a)+\frac{\sinh^2 a-1}{\cosh^3 a}\cdot\frac{(\tilde{\varphi}-a)^2}{2}+ \tag{3.4a}$$
$$+\frac{5\sinh a-\sinh^3 a}{\cosh^4 a}\cdot\frac{(\tilde{\varphi}-a)^3}{6}+\cdots.$$

Thus, if $a=0$ (when α^* is small — smaller than 1 — the high injection level) we obtain:

$$I_3^*=\frac{1}{2}-\frac{17}{140}\alpha^{*2}+\frac{57\cdot 58}{1536}\alpha^{*4}-\cdots,\quad \alpha^*\leqslant 1, \tag{3.4b}$$

and if $a=2.095$ we obtain:

$$I_3^*=0.63924-0.26658\,\alpha^*+0.0488\,\alpha^{*2}-0.0035\alpha^{*3}+ \tag{3.4c}$$
$$+0.000221\,\alpha^{*4}-\cdots,\quad 1\leqslant\alpha^*\leqslant 3;$$

— the second way is the *graphical* intergration.

The results of the calculation in both the above ways are given in Chapter 4. (The departures in relation to [22] are seen to *be very great:* they may be even above 200%).

The expression

$$I_3\approx\frac{2}{3\alpha},\quad(\alpha\approx\alpha^*), \tag{3.4d}$$

is used for large α's in [22], but we have shown in [21] that the approximation:

$$I_3^*(\alpha^*)\approx\frac{2}{3\alpha^*}\left[1-\frac{\sqrt{3}}{2}e^{\frac{\pi}{2}\alpha^*}\left(1-\frac{1}{3\sqrt{3}\,\alpha^*}\right)\right],\quad(\alpha^*>4) \tag{3.4e}$$

is slightly better.

4° — The integral I_3^* for the abrupt asymmetric transition is *not treated* at all (in [22]). One may show that the definition (3.2c) is still valid for the *very asymmetric* transition ($P\gg N$), but the eqn. (3.2b) is *not valid* for $\tilde{\varphi}$ and α^*, but

$$\tilde{\varphi}\approx-2\alpha^*(t-1)^2,\quad \alpha^*\approx\frac{\alpha}{2}+\frac{1}{4}\ln\frac{P}{N},\quad\left(t=\frac{x}{x_N^*}\right). \tag{3.5a}$$

The following two ways are treated here:

— *graphical* integration

— and the representation of $\operatorname{sech}\tilde{\varphi}$ as the sum*) of the infinite geometric progression of the factor $q=-e^{2\tilde{\varphi}}$ (since $\tilde{\varphi}$ is always negative, see also Fig. 2.3) so that

$$I_3^*=\sum_{n=1}^{\infty}(-1)^{n-1}\sqrt{\frac{\pi}{8\alpha^*(2n-1)}}\cdot\Phi_n(x_n)=$$

$$=\sqrt{\frac{\pi}{8\alpha^*}}\left[\sum_1^4\frac{(-1)^{n-1}}{\sqrt{2n-1}}\cdot\Phi_n(x_n)+\sum_5^{\infty}\frac{(-1)^{n-1}}{\sqrt{2n-1}}\right], \tag{3.5b}$$

$$\Phi_n=\frac{2}{\sqrt{2\pi}}\int_0^{x_n}e^{-z^2/2}dz,\quad x_n=2\sqrt{\alpha^*(2n-1)},$$

is obtained, with the fact that $\Phi_n\approx 1$ for $n\geqslant 4$, being used in the final expression for I_3^*. Thus this problem consists of finding tabulated values for Φ_n ($n=1$, 2, 3, and 4) and the sum $\sum_5^{\infty}\cdots$; the last one is slowly convergent, but it might be expressed by means of Riemann's ζ-function ([21]) with sufficient accuracy. Finally

$$I_3^*(\alpha^*)=\frac{0.6267}{\sqrt{\alpha^*}}\left[0.1756+\sum_1^4\frac{(-1)^{n-1}}{\sqrt{2n-1}}\cdot\Phi_n(x_n)\right], \tag{3.5c}$$

($\alpha^*\geqslant 0.4$), is obtained.

The value of the *first derivative* $dI_3^*/d\alpha^*$ is also of the great importance. Starting from (3.5a) for $\tilde{\varphi}$ and (3.2c)

$$\frac{dI_3^*}{d\alpha^*}=\frac{d}{d\alpha^*}\left(\frac{1}{\sqrt{32\alpha^*}}\int_0^{2\alpha^*}\frac{dy}{\sqrt{y}\cosh y}\right)=\cdots=\frac{1}{2\alpha^*}\left(I_3^*-\frac{1}{2\cosh 2\alpha^*}\right) \tag{3.5d}$$

is obtained, which is *quite exact.*

The results of the calculations based on eqns. (3.5c and d) will be given in Chapt. 4.

b. — Diffusion currents — Higher injection level

1° — If the space charge and the electric field effect (drift components) are neglected in the vicinity of $x_{N;P}^*$, then eqns. (2.2) and (2.3) in the steady state become purely „*diffusive*" for *minority* carriers:

$$\frac{d^2n_P}{dx^2}=\frac{\delta n_P}{L_{nP}^2}\equiv\frac{n_P-n_{0P}}{L_{nP}^2},\quad x\leqslant x_P^*,\quad \frac{d^2p_N}{dx^2}=\cdots. \tag{3.6a}$$

The boundary conditions must be retained in their full form (2.16), since the value of χ is not much smaller than unity ($\overset{\circ}{n}_P\neq N\chi$, $\overset{\circ}{n}_N\neq\overset{\circ}{n}_{oN}$ etc):

*) At the suggestion of Ing. M. Smiljanić, Inst. of Physics— Beograd.

it is implied that $p_N^{\circ} \approx$ or $\gg N$, and thus if „only" $p_N^{\circ} = \frac{N}{2} = \frac{P}{2} = n_P^{\circ}$ for the symmetrical transition, then χ is already about 0.75.

2° — Thus, for the *medium* levels $\left(p_N^{\circ} \approx \frac{N}{5},\ \chi \approx \frac{1}{4}\right)$ for the *symmetrical linear transition* according to (3.6a), ([21]) is obtained:

$$J_d = e\left[\frac{D_{nP}(n_P^{\circ} - n_{oP}^{\circ})}{L_{nP} + x_N \cdot I_{2n}(U^*)} + \frac{D_{pN}(p_N^{\circ} - p_{oN}^{\circ})}{L_{pN} + x_N \cdot I_{2p}(U^*)}\right], \tag{3.6b}$$

where the above index „0" denotes the point $x = x_{N,P}^*$ and where the *integrals* I_2 have also the same significance like those in (3.2b). Actually, (3.6b) represents the „corrected" well-known relation for the low level of the injection, as, for instance (2.29c) with $J_s = const$. The general treatment of this problem will be given in the point *c*.

Supposing that $L_{nP;pN} \gg x_N I_2$ („thin" transitions), we obtain from (3.6b) for *all kinds* of transitions and medium levels:

$$\begin{aligned} J_d = \cdots = en_i^2 \Bigg\{ & \frac{D_{nP}}{L_{nP} P} \cdot \frac{e^{e\tilde{U}'} \cdot \left[0.5 + \left(\frac{1}{4} + \frac{N}{P}\chi\right)^{1/2}\right]^{-1} - 1}{e^{e\tilde{U}'} - 1} + \\ & + \frac{D_{pN}}{L_{pN} N} \cdot \frac{e^{e\tilde{U}'} \cdot \left[0.5 + \left(\frac{1}{4} + \frac{P}{N}\chi\right)^{1/2}\right]^{-1} - 1}{e^{e\tilde{U}'} - 1} \Bigg\} (e^{e\tilde{U}'} - 1) = \\ & = \frac{en_i^2}{N}\left[\frac{D_{nP}}{L_{nP}} + \frac{D_{pN}}{L_{pN}}\right] \cdot \left[\frac{e^{e\tilde{U}'}}{\frac{1}{2} + \left(\frac{1}{4} + \chi\right)^{1/2}} - 1\right], \end{aligned} \tag{3.6c}$$

where the final expression is valid only for the symmetrical transition (abrupt or graded). Eqn. (3.6c) is implied to be valid also for the *lower* levels and it transforms into the expressions of the type (2.29c) in the limit.

3° — Eqn. (3.6a) would be approximately valid further on for *high levels*, but not only $p_N^{\circ} \gg N$ or $n_P^{\circ} \gg P$ but also $n_N \gg N$ or $p_P \gg P$ is valid for *majority* carriers. Therefore, *almost the same large* diffusion currents of the majority carriers appear, and in order to keep quasi-neutrality, a field that should "almost" cancel these currents, must be created.

Regarding the *asymmetrical* transition $\left(\frac{P}{N} \gg 1\right)$ in the limit $(\chi \to 1,\ U' \to U_D)$, it can be easily shown that the diffusion „constant" for N-region is:

$$D_p(p) = D_{p0}\left(1 + \frac{p}{p+N}\right),\ (p = p_N \gg N), \tag{3.7a}$$

i.e. almost $2 \cdot D_{p0}$, while $n_P \approx P \approx p_P$ and $D_{nP} = D_{n0} = const$. We should emphasize that this is *not valid* for the symmetrical transition, because, according to (2.16c) the condition $P_N \gg N$ cannot be attained.

Consequently, (3.6b and c) is *still valid*, and it is also valid for the asymmetrical transition ($P \gg N$) with $D_{pN} \to D_p(p)$ according to (3.7a) and

$$L_{pN} \to L_p(p) = L_{p0}\left(1 + \frac{p}{p+N}\right)^{1/2}, \quad (p = p^\circ). \tag{3.7b}$$

c. — Dependences $J(U)$

1°—Regarding the medium *injection levels*, considering that $\sinh \frac{e\tilde{U}'}{2} \approx \frac{\exp \frac{e\tilde{U}'}{2}}{2}$, according to (3.2a) and (3.6c), since $\exp e\tilde{U}' \gg 1$ and $\exp e\tilde{U}' [\cdots]^{-1}$ in (3.6c) are much larger than 1,

— for the *asymmetric* transition with $P \gg N$, $\frac{N}{P}\chi \ll 1$ and $\frac{P}{N}\chi \gg 1$, we obtain:

$$J_{dir} = A_{g-r} \cdot B_{g-r} \cdot e^{e\tilde{U}'/2} + J_{sn} B_{dn} e^{e\tilde{U}'} + A_d B_{dp} e^{e\tilde{U}'/2}, \tag{3.8a}$$

where A's and J_{sn} are the constants and B's are quasi-constants:

$$A_{g-r} = \frac{2en_i}{\tau}; \quad B_{g-r} = I_3^*(U^*) \cdot x_N; \quad J_{sn} = \frac{en_i^2 D_{nP}}{L_{nP} \cdot P};$$

$$B_{dn} = \frac{2}{1 + \left(1 + \frac{4N}{P}\chi\right)^{1/2}} \approx C^{te} \approx 1; \quad B_{dp} \approx 1; \tag{3.8b}$$

$$A_d = \left(en_i^2 \frac{D_{pN}}{L_{pN} N}\right) \frac{N}{n_i} \equiv J_{sp} \cdot \frac{N}{n_i} = \frac{en_i D_{pN}}{L_{pN}},$$

— but the following is obtained for the *symmetric* transition ($P = N$) for the diffusion current:

$$J_d = J_s \frac{PN}{n_i^2}\left(\sqrt{1+4\chi} - 1\right), \quad \left(\chi = \frac{n_i^2}{PN} e^{e\tilde{U}'}\right). \tag{3.8c}$$

It can be easily shown that the ratio of the second and the third term in (3.8a) is always smaller that $(N/P)^{1/2}$ for $U' \leqslant U_D$ (the medium level) i.e. regarding the two ,,diffusion'' components, the component $\sim e^{e\tilde{U}'/2}$*) *prevails* in the asymmetrical transition and the total current has two components both directly proportional to the $\exp \frac{e\tilde{U}'}{2}$.

We can see from (3.8c) that the further expansion of the expression $(1+4\chi)^{1/2}$ is *useless*, because $4\chi \gg 1$ can never be for $U' \leqslant U_D$, i.e. the component $\sim \exp. \frac{e\tilde{U}'}{2}$ *does not exist* here in its real sense. This fact is not noticed to be emphasized by the others.

*) Evidently, the component $\sim \exp e\tilde{U}'$ prevails for $U' > U_D$, but those are already higher levels; apart from that, U' is smaller than the applied voltage U for the voltage drop beyond the transition, thus the condition $U' > U_D$ can hardly be attained in practice.

2° — For *high levels* the diffusion current is obtained for the *asymmetric* transition according to (3.7a) and (3.7b):

$$J_d = J_{pd}(x_N^*) + J_{nd}(x_P^*) \approx J_{pd}(x_N^*) = \cdots \approx$$

(3.8d)

$$\approx \frac{D_{pN}(1 + p_N^\circ/p_N^\circ + N)^{1/2}}{L_{PN}} (p_N^\circ - p_{0N}^\circ) \sim p_N^\circ .$$

We see that $J_d \sim p_N^\circ \sim \sqrt{\chi} \sim \exp\frac{e\tilde{U}'}{2}$ according to (2.16a). Let us remark that the above fact in [12] (page 159) is taken from otherwise correct conclusion that $D_p \to 2D_{pN}$ for high levels, but the interrelation between this conclusion and the fact mentioned above is, in our opinion, not clear.

3° — As a *conclusion*, by forming the following relations according to the above:

(3.8e)
$$\gamma_{rd}(U) = \frac{J_{g-r}}{J_d}$$

it is possible to show ([21]) that this ratio between recombination and "diffusion" currents *constantly decreases* in *any* transition (abrupt or graded, symmetrical or asymmetrical) when the polarization (for $U' \leqslant U_D$) is direct, with the increase of U, which is obvious even physically, but according to the conception which does not take the band edges correction into consideration (does not distinguish α^* from α according to the eqn. 3.2b) the correct result is *not obtained*. Thus, the relations (3.5c) and (3.5d) have also a great principal significance.

Finally, by forming the ratio $\xi = \gamma_{rd}(U'=0)/\gamma_{rd}(U'=U_D)$ we obtain approximately:

(3.8f)
$$\xi_{sym} \approx \frac{1.11\,N}{\left(\ln\frac{N}{n_i}\right)^{2/3} \cdot n_i} \quad \text{and} \quad \xi_{as} \approx \sqrt{2} + \frac{D_{nP} L_{pN}}{D_{pN} L_{nP}} \sqrt{\frac{N}{P}},$$

for the linear symmetrical and abrupt asymmetrical ($P \gg N$) transition, respectively. We see that ξ_{sym} may be *very large* $\left(\xi_{sym}^{min} \approx 2{,}9 \text{ for } \frac{N}{n_i} = 1.95\right)$ which is *not the case* with ξ_{as}, because $D_{nP} L_{pN} \sqrt{\frac{N}{P}} \Big/ D_{pN} L_{np}$ is never larger than unity ´ 'en not for In Sb where $D_n/D_p = \mu_n/\mu_p \approx 100$). Hence, with the *increase of* nsition *asymmetry*, the ratio γ_{rd} remains *„less suitable"* if the injection el is increased: the participation of the recombination currents from the nsition decreases less.

d. — Higher levels — General relations

1° — Let us be satisfied with the solution for J_{g-r} from the point *a*, try to find the more correct — general solutions for the *„diffusion"* ents beyond the transition $(x \leqslant x_P^*,\ x \geqslant x_N^*)$ taking into consideration the *'tions* described in the paragraph 3° — point *b*. By using the model and

all the transport relations from (2.2) to (2.7), except Poisson's (2.4) instead of which the *quasi-neutrality* condition*) is written

(3.9a) $$n_N(x) \approx p_N(x) + N(x), \quad x \geqslant x_N^*;$$

$$p_p(x) \approx n_p(x) + P(x), \quad x \leqslant x_P^*,$$

emphasizing that both N and P *may depend* upon x (graded transition) for the one-dimensional and planar case, when according to (2.3), with $r_{ne} = r_{pe} = r_e$, we get:

(3.9b) $$\frac{\partial}{\partial x}(J_n + J_p) \equiv \frac{\partial}{\partial x} J = 0 \quad \therefore \quad J = const.$$

i.e. the *total current constant*, for the field one can easily obtain from (2.3):

(3.9c) $$K(x) = \frac{J - e\left(D_n \frac{\partial n}{\partial x} - D_p \frac{\partial p}{\partial x}\right)}{e\,[n(x) \cdot \mu_n + p(x) \cdot \mu_p]}, \quad (\mu,\ D = C^{te}).$$

By noting the derivatives with respect to x with „prim" and „second", after long but simple calculation, the *general equation* for the holes in the N-region ($x \geqslant x_N^*$) is obtained from (3.9c), (3.9a) and (2.2) together with (2.7):

(3.9d) $$\underset{\text{(I)}}{p''} + \underset{\text{(II)}}{(b-1) \cdot Q(p, N) \cdot (p')^2} - \underset{\text{(III)}}{\frac{J}{eD_{pN}} Q(p, N) \cdot Q_1 \cdot p'} +$$

$$+ \underset{\text{(IV)}}{Q(p, N) \cdot Q_2 \cdot p} - \underset{\text{(V)}}{\frac{R(p, N)}{D_{pN} \tau_0 b}} = 0,$$

where

$$Q_1 = 1 + \frac{eD_{pN}}{J}\left(\frac{b-1}{N} p - b\right) \cdot N', \quad Q_2 = \frac{F \cdot N'' + N'\left(\frac{J}{eD_{pN}} - bN'\right)}{N};$$

$$Q = \frac{N}{F(p, N) \cdot (2p + N)}, \quad F(p, N) = p(b+1) + bN;$$

(3.9e) $$R(p, N) = \tau_0 \frac{r_e \cdot F}{2p + N}, \quad r_e \equiv \frac{\delta p}{\tau_0} \cdot \frac{1 + c\frac{\delta p}{N}}{1 + a\,\delta p/N}, \quad c = \frac{N}{N + 2p_0},$$

$$a = \frac{\tau_\infty N}{\tau_0 (N + 2p_0)}; \quad \tau_\infty = \tau_{n0} + \tau_{p0}, \quad \tau_0 = \frac{p_0 + B_v R}{N + 2p_0} \tau_{n0} + \frac{n_0 + B_c R}{N + 2p_0} \tau_{p0},$$

$$b = \mu_n / \mu_p; \quad \tau_{p0} = \frac{1}{\gamma_p N_R}, \quad \tau_{n0} = \frac{1}{\gamma_n N_R}; \quad \delta p = p - p_0.$$

*) Actually, (3.9a) should be considered as Poisson's relation with $\rho \approx 0$; consequently, it is always necessary. Nevertheless, $dK/dx = 0$ and $K = const.$ would result from $\rho = 0$, which is too rough, so that (3.9c) is used for K.

The equation of the type (3.9 d) was obtained even earlier (e.g. in [34] to [37]) but this one is more *general* than those given up to now, because it contains both N' and N'', which is important for graded transitions, at least on principle.

Only the abrupt transition will be treated here ($Q_1 = 1$ and $Q_2 = 0$), but even then (3.9 d) has several advantages:

— only the direct recombination (formally with $a = 0$ in the expression 3.9 e for r_e) is treated in [34], but with us ($a \neq 0$) both are possible: the direct and the indirect recombination according to the model (2.7); apart from that, the case of large p's is treated in [34] only for $J \approx 0$, but actually the contrary would be more logical.

— Eqn. (3.9 d) has also been attained in [35] (without Q_1 and Q_2) but in actual solution it is considered that $\tau_p = const.$, hence even rougher than in [34]; apart from that, the terms (II) and (III) are neglected in [35] and some restrictions are not documented.

— The problem is presented in [37] more widely than with (3.9 d): the impurities are not completely ionized, $\varrho \neq 0$, but in actual solution it is considered that $\varrho \approx 0$, and the differential equations are linearized without special explanation*); apart from that, the conception of external field (homogeneous) is applied in [37], but in our opinion, it would be more correct to consider external voltages.

2° — Regarding the *abrupt transition* ($Q_1 = 1$ and $Q_2 = 0$) by introducing the symbols:

$$y = p/N, \quad y' = \frac{dy}{dx} = p'/N = z, \quad y'' = z\frac{dz}{dy};$$

(3.10a)
$$G = L_{pN0} \cdot z, \quad (L_{pN0}^2 = D_{pN0}\tau_0)$$
$$A = \frac{J L_{pN0}}{e D_{pN} N b},$$

the equation (3.9 d) under the condition that $\delta p \approx p$ (*higher* levels), is obtained in the form:

(3.10b)
$$G \cdot \frac{dG}{dy} + N(G, y) = 0$$

with

$$N(G, y) = Q'(y) \cdot (\beta G^2 - GA) - \gamma(y),$$

(3.10c)
$$\beta = \frac{b-1}{b}, \quad B = \frac{b+1}{b}, \quad Q' = \frac{1}{(1+By)(1+2y)},$$
$$\gamma = \frac{1+By}{1+2y} \cdot y \cdot \frac{1+cy}{1+ay}.$$

Hence we have the *generalized*: concentration gradient G, concentration y and current A. We remark that it is *sufficient* to know the dependence $G(y)$ i.e. $p'(p)$ and not necessarily $p(x)$, for our problem.

*) The detailed comparison of all the terms with (I) is carried out in [21], and (3.9 d) is shown not to be a linear equation on principle, (let alone the eqn. with constant coefficients), and actually it must be solved *complete*.

The equation (3.10b) is Abel's differential equation for which, as far as we know, there is no general solution in quadratures. It exists only for $A=0$ and it is as follows:

$$G\Big|_{A,\,J=0} = -\left(\frac{2I_y}{S^2}\right)^{1/2},\quad I_y \equiv I(y) = \int_0^y S^2\cdot\gamma(y)\,dy,\quad S(y) = \frac{1+2y}{1+By}, \tag{3.11a}$$

which can be easily obtained by the method of integrating factor (independent of G).

By using the integrating factor which depends on G (which is not adequate mathematically), the *approximate solution* (explained in detail in [21]) in the form

$$\boxed{(3.11b)}\qquad -G = \left(\frac{2I_y}{S^2} + H^2 A_1\right)^{1/2} - HA_1,$$
$$A_1 = A/4\beta,\quad H(y) = 1 - \frac{1}{S^2}$$

is obtained.

The value of the proposed solution will be checked in Chapter 4, but now it is only clear that it coincides with (3.11a) for $A=0$, and that it satisfies the boundary condition $G=0$ for $y=0$. We remark that the integral I_y has a solution expressible by *elementary functions* and that (3.11b) is *roughly* valid both for *graded* transitions and for *all* injection levels.

It is valid for *very low* levels ($y\to 0$), since then

$$S\approx 1,\quad \gamma\approx y,\quad I_y\approx y^2/2,\quad H\approx 2\beta y, \tag{3.11c}$$

and from (3.11b)

$$-G\equiv -L_{pN0}\,z = \left(\sqrt{1+A^2/4} - A/2\right)\cdot y, \tag{3.11d}$$

which completely coincides with the solution of the equation

$$y'' - \frac{A}{L}y' - y/L^2 = 0 \tag{3.11e}$$

obtained from (3.9d) for $y\to 0$ and $|y'|\gg(y')^2$, which is usually used for low levels (and abrupt transitions — $Q_2=0$).

3° — With $c=1$, for the *very high levels* ($y\to\infty$) according to (3.11b), the following solutions are obtained

$$\boxed{(3.11f)}\qquad -G = \left[\frac{b+1}{2ab}y^2 + \frac{b+1}{a^2b}(a-1)\,y\right]^{1/2} \approx \left(\frac{b+1}{2ab}\right)^{1/2}\cdot y \quad \text{for } a\neq 0,$$
$$-G = \left[\frac{b+1}{3b}y^3 + \frac{3b+1}{4b}y^2 + \frac{b-1}{4b}y\right]^{1/2} \approx \left(\frac{b+1}{3b}\right)^{1/2}\cdot y^{3/2} \quad \text{for } a=0,$$

i.e. *different* for the indirect ($a\neq 0$) and direct ($a=0$) recombination and *more general* than the corresponding relations in [34] and [37]. The same result is obtained also from the immediate solution of (3.9d), when $p\to\infty$ and when the terms II, III and IV are neglected.

We see that G does not depend upon the current (parameter A) for very high injection level.

By using (3.7a) according to (3.6b) the diffusion current is obtained at *high* levels for the *asymmetrical* transition:

$$J_d \approx J_N^\circ = -\frac{2D_{pN}\cdot N}{L_{pN0}+x_N^*\cdot I_{2p}(\alpha^*)}\,G^\circ,\quad G^\circ = G(y^\circ), \tag{3.11g}$$

with G° according to (3.11f) and $y^\circ \equiv p_N^\circ/N$ according to (2.16a). We *conclude* from (3.11g) and (3.11f) that J_{dir}

— contains the quasi-constant term besides the terms according to (3.8a) for $a \neq 0$:

$$A_{as}^\infty = \frac{2eD_{pN}\cdot N}{L_{pN0}}\left[\frac{(a-1)^2(b+1)}{2ba^3}\right]^{1/2}, \tag{3.11h}$$

since $\sqrt{x^2+2x} \approx x+1$ for $x \gg 1$.

— but for the *direct* recombination ($a=0$) the following term will appear instead of the term $Ad\ B_{dp}\,e^{e\tilde{U}'/2}$ in (3.8a):

$$A_{as_0}^\infty \cdot \left[\left(\frac{n_i}{N}\right)^{3/2} e^{\frac{3}{4}e\tilde{U}'} + \frac{3b+1}{b+1}\cdot\frac{3}{8}\left(\frac{n_i}{N}\right)^{1/2} e^{e\tilde{U}'/4} + \frac{b-1}{b+1}\cdot\frac{3}{8}\left(\frac{N}{n_i}\right)^{1/2} e^{-e\tilde{U}'/4}\right], \tag{3.11i}$$

$$A_{as_0}^\infty = \frac{2eN D_{pN}}{L_{pN0}}\sqrt{\frac{b+1}{3b}},$$

i.e. the exponents $\frac{3}{4}e\tilde{U}'$ and $\frac{e\tilde{U}'}{4}$ which are *different* from those for $a \neq 0$. The results expressed with (3.11h) and (3.11i) are very interesting, and as far as we know, they are new; however, the checking of their practical significance is beyond the frame of this paper.

3.2. — Limiting regimes at reverse polarization

a. $\mu(K)$-dependence — hot electrons

1° — When the field is increased, the carrier distribution ceases to be equilibrium: the carriers do not „have time" to transfer all their energy to the lattice and they become *hot*. The notion of electron temperature $T_e > T_L = T$ is introduced in ([24] — 1951). Mobility begins to depend on the field — it decreases. Without entering into the explanation of this phenomenon*), let us emphasize that the essential role is played here by the scattering by the *optical phonons* (of energy $\hbar\omega_0$) and that the *saturation* effect of the drift velocity appears there. This velocity becomes

$$v_d \equiv \mu K = v_s \approx \sqrt{\frac{\hbar\omega_0}{2m^*}} \quad \therefore\ \mu \sim 1/K \tag{3.12a}$$

for the field K' of the order 10^3 V/cm in the wide temperature region (50 to 400°K for Ge); v_s are of the order 70 km/s. K' is the field in the „middle" of the borderline region $\mu = c^{te} - \mu \sim 1/K$.

*) It seems that the first such theory is given in 1934 (L. Landau, A. Kompanejez. Phys. Z. Sowjet., 6, 163) and later on it was treated by many authors ([7] — 1937; Seitz; Conwell — [30] — 1952; Gunn, and many others).

Specific conditions mentioned at the beginning of this Chapter prevail in P-N transition, so that the following questions must be cleared up:

— Is the transition width (Λ_ρ) small in relation to the length of the free path of the carriers ($\Lambda_{n,p}$), so that the complex effects appear as if being in thin layers (see for instance [15])? However it is possible to prove that:

$$\Lambda_\rho \gg \Lambda_N,\ \Lambda_p; \tag{3.12b}$$

— The „nakedness" of the impurity may cause the *quantitative* change of the scattering character on the ionized impurities i.e. the carriers may not be of „band" but of „*quasi-polaron*" character; it was shown in [21] in detail that the carriers in the atomic semiconductors *may*, however, *be treated* as „band" carriers and that for the mobility a *more suitable* expression is that due to Conwell [30], which is usually considered as less accurate than that given in [29];

— Relaxation time, or rather distribution function may depend on x explicitly, and not only implicitly throughout $K=K(x)$; it was shown by Čujenkov ([31] — II) that $\tau \neq f(x)$ if

$$|K(x)| \gg \left|\frac{dK}{dx}\right| \frac{\bar{\Lambda}_n}{\sqrt{\eta_0}},\quad \eta_0 \equiv \eta_{0pt} = \frac{3\hbar\omega_0}{4C_0 W^*}, \tag{3.12c}$$

where η_0 is a relative energy given by the carrier of energy $W=W^*$ to the optical phonons, W^* — the threshold of the atomic ionization energy, and

$$C_0 = \coth \frac{\hbar\omega_0}{2kT}. \tag{3.12d}$$

Eqn. (3.12c) may be shown, what is done in [21], to *satisfy any graded transition*, but regarding an abrupt transition — only if $\Lambda_0 \gg \bar{\Lambda}_n/\sqrt{\eta_0}$ holds, which *need not always* be the case.

Hence, nothing changes in the machanism of scattering processes in relation to the machanism in the bulk, under the condition that (3.12d) is fulfilled.

2° — The common effect of the optic and acoustic phonons were treated by Bok [33] in an alternate way (strong interaction of carriers) and under the condition that

$$m^* v_\Phi \ll kT \quad \text{and} \quad \frac{m^* v_d^2}{2} \ll kT_e, \tag{3.13a}$$

where v_Φ is the phonon velocity and v_d — the drift velocity. He got the following parametric dependence for v_d and μ:

$$\frac{K}{v_{ds}} = \frac{1}{\mu_s} = F_s(\xi_e) = \xi_e/\mu_{s0},\quad \xi_e = \left(\frac{T_e}{T}\right)^{1/2} \tag{3.13b}$$

$$K v_{ds} = \mu_s K^2 = G_s(\xi_e), \tag{3.13c}$$

where the subscript „s" denotes the type of scattering („*ac*" — acoustical and „*opt*" — optical). As F_s's and G_s's are the known functions (different

for various s's) — μ_s's, v_{ds}'s and T_e for the given field (and T) can be obtained from (3.13). The *resultant* mobility μ_Φ can be obtained according to:

$$\frac{1}{\mu_\Phi}=\frac{1}{\mu_{ac}}+\frac{1}{\mu_{o\,pt}}=F_{ac}+F_{opt}=\xi_e\left(\frac{1}{\mu_{aco}}+\frac{1}{\mu_{opto}}\right) \tag{3.13d}$$

$$Kv_{d\,\Phi}=K\left(\frac{1}{v_{dac}}+\frac{1}{v_{d\,opt}}\right)^{-1}=\mu_\Phi K^2 \therefore \frac{1}{\mu_\Phi}=K^2\left(\frac{1}{G_{ac}}+\frac{1}{G_{0\,pt}}\right). \tag{3.13e}$$

According to ([38])

$$G_{ac}=\frac{2v_\Phi^2}{\mu_{aco}}\xi_e^2\left(1-\frac{1}{\xi_e^2}\right) \text{ and } G_{opt}=\frac{\hbar\omega_0}{C_0\,\mu_{opto}\cdot m^*}\xi_e, \tag{3.13f}$$

we obtain from (3.13d) and (3.13e) for the larger field ($\xi_e^2\gg 1$)

$$\left(\frac{\mu_\Phi}{\mu_{\Phi_0}}\right)^2=\left\{\left[\frac{C_0 m^* v_\Phi{}^2}{\hbar\omega_0}\cdot\frac{\mu_{opto}}{\mu_{aco}}\right]^2+\frac{1}{K^2}\left(\frac{2v_\Phi{}^2}{\mu_{\Phi 0}\,\mu_{aco}}\right)\right\}^{1/2}- \tag{3.13g}$$

$$-\left[\frac{C_0 m^* v_\Phi{}^2}{\hbar\omega_0}\cdot\frac{\mu_{opto}}{\mu_{aco}}\right],\quad \frac{1}{\mu_{\Phi_0}}=\frac{1}{\mu_{aco}}+\frac{1}{\mu_{opto}}.$$

The member $(\ldots)/K^2$ is $\ll[\cdots]^2$ for the fields large enough and by developing $\{\cdots\}^{1/2}$ into the series, we easily get

$$\left(\frac{\mu_\Phi}{\mu_{\Phi_0}}\right)^2\approx\frac{\hbar\omega_0\,\mu_{\Phi_0}}{C_0 m^*\mu_{opto}\,\mu_{\Phi 0}{}^2}\cdot\frac{1}{K^2}\therefore v_s=\mu_\Phi K=\sqrt{\frac{\hbar\omega_0}{C_0 m^*}}\left(\frac{\mu_{aco}}{\mu_{aco}+\mu_{opto}}\right)^{1/2} \tag{3.13h}$$

which is at least semi-quantitatively correct, and it agrees with the results of other authors mentioned earlier. The value of this method is in the fact that the explicit dependence $\mu_\Phi(K)$ is obtained-according to (3.13g), and as $(\mu_{\Phi_0}/\mu_\Phi)^2=\frac{Te}{T}=\xi_e^2$ we have also the *dependence* $T_e(K)$.

We remark that the dependence $\mu_\Phi(K)$ is incorrectly deduced in [33] (eqn. 3.13g is not obtained) because the ,,energetic" expression (3.13e) according to (3.13c) is noted in an non-adequate way:

$$\mu_\Phi K^2=G_{ac}+G_{opt},\ ([33]\text{ — p. 128}), \tag{3.13i}$$

and there is even an unadequate conclusion that the saturation exists at lower ξ_e's i.e. K's (because the member with ξ_e prevails in 3.13i for the lower K) and that there is a further (,,acoustic") increase of $v_d\sim\sqrt{K}$ after the ,,Ohmic" part of $v_d\sim K$ and the part od $v_d=v_s=const.$

b. — Avalanche ionization

1° — According to Čujenkov [31] — II, the distribution function in the SC-region is shown not to depend explicitly upon x, i.e. avalanche ionization phenomenon may be treated in the same way as in the homogeneous semiconductor, if the condition (3.12c) is fulfilled.

We should note that the avalanche ionization in solids may be equalled to impact ionization in gases (glowing discharge; Townsend, as early as in 1901) only formally, because in reality there are some *essential differences:*

— no pair- carriers are created in gases: hardly mobile ions ($\mu_I \ll \mu_n$, but $\mu_n \approx \mu_p$) are there instead of holes, and the conductance is *monopolar*;

— drift velocity has *the saturation* long before the avalanche effect, as it was described in the Section **a.** — so that the *direct acceleration* in the field is almost prevented and the impact ionization takes place *indirectly*: the field increases the average thermal velocity (v_T) i.e. carrier temperature, and at *such distribution* the fastest carriers ionize neutral atoms which have the *ionization threshold* ([31])

$$W^*_{n,p} \equiv \frac{m_{n,p}\,(v^*_{n,p})^2}{2} = W_I \frac{1+2\dfrac{m_{n,p}}{m_{p,n}}}{1+\dfrac{m_{n,p}}{m_{p,n}}} \tag{3.14a}$$

and ionization probability

$$w_I(\eta) \approx w_0 \sqrt{\frac{\eta-1}{\eta}};\ \eta = W/W^*,\ \left(W = \frac{\hbar^2 k^2}{2m^*}\right), \tag{3.14b}$$

where $W_I = W_g$ — if the intrinsic ionization is concerned, and $w_0 = \text{const.}$ — *ionization parameter.*

In [31], for the field K' defined at (3.12a), the following is obtained:

$$K' \approx 1.13\ \frac{\sqrt{m^*}\,C_l^2 \cdot \hbar\,\omega_0 \cdot v_\Phi^2}{\mu_{aco}\,(kT)^{3/2}}\,\zeta_\Phi^2 \tag{3.14c}$$

where ζ_Φ is the parameter of the ratio between scattering by optical and that by acoustical phonons (for Ge and Si of N-type, it is 2 to 2.5 for the wide temperature range); all other signs are as before.

2° — By introducing the characteristic lengths of the free path (Λ_0), the field (K_0) and the *ionization* probability *parameter**):

$$\Lambda_0 = \Lambda_{ac}\,\frac{\eta_{opt}}{\eta_{ac}} \cdot \frac{1}{\zeta_\Phi^2} \geqslant \frac{\hbar\,\omega_0}{2C_0 W^*}\,[\mu\,m] \approx 50\ \text{Å},\ \eta_{ac} = \frac{2mv_\Phi^2}{kT}, \tag{3.15a}$$

$$K_0 = \frac{W^*}{e\,\Lambda_0}\sqrt{\eta_{opt}} \approx 10^4 - 10^5\ \text{V/cm},\ K^* = \frac{W^*}{e\,\Lambda_0} \tag{3.15b}$$

$$p_0 = w_0\,\frac{\Lambda_0}{v^*\,\eta_{opt}} \approx w_0 \cdot 10^{-12} = 1\ \text{to}\ 10^4,\ \left(v^* = \sqrt{\frac{2W^*}{m^*}},\ \eta_{opt} = \frac{3\hbar\,\omega_0}{4C_0 W^*}\right), \tag{3.15c}$$

Čujenkov obtains the expression for the *coefficients* ($\alpha_{n,p}$) *o tion*, for stronger fields, in *which only* we take an interest (z

$$\alpha(z) = \frac{3}{\sqrt{\pi}}\,\frac{v^*\,\eta_{opt}}{\Lambda_0 v_s} \cdot z\,e^{-z^2}\left[\frac{2 \cdot 3^{1/4} \cdot \Gamma(3/4)}{3 \cdot \Gamma(1/4)} \cdot \frac{p_0^{1/4}}{z^{3/2}} - \right. \tag{3.16a}$$

*) Numerical values in (3.15) are calculated in [21] according to for $T = 300°\text{K}$; w_0 exists within the range of 10^{12}s^{-1} in case of week ioni lattice scattering, and up to 10^{16}s^{-1} when the ionization is preponderant.

which is very similar to that of Kel'dyš [38], where the attention is paid*) to whether the dielectric constant (ε) is near the unity or $\varepsilon \gg 1$ (which would correspond to Ge, Si, GaP and GaAs):

$$\alpha(z) = z^{\nu} e^{-z^2} \left[2 \cdot 3^{1/4} \frac{\Gamma(3/4)}{\Gamma(1/4)} \cdot \frac{(p_0')^{1/4}}{z^{3/2}} - \frac{1}{2} \right] \cdot C(z), \tag{3.16a'}$$

where $C(z)$ almost do not depend on z, and where $\nu \approx 2$.

The compositions of eqns. (3.16a) and (3.16a') are similar, but we have decided to take (3.16a), because (3.16a') cannot be used for the limiting case $K \to \infty$ ($z \to 0$), then $\alpha \sim z^{1/2}$, but if K increases, then ionization must also increase.

By introducing the notation

$$\Lambda^* = \Lambda^\circ / k_x, \quad x = \frac{W^*}{eK\Lambda^*} = z\sqrt{\frac{A}{\psi}} \sim \frac{1}{K}, \quad \psi = k_\psi \frac{\hbar\omega_0}{C_0 W^*},$$
$$A = \frac{4}{3} k_x^2 k_\psi, \quad B = 2\frac{\Gamma(3/4)}{\Gamma(1/4)} \left(w_0 \frac{3\Lambda^*}{v^* k_x} \right)^{1/4}, \quad K_{n,p} \equiv \frac{W^*_{n,p}}{e\Lambda^*_{n,p}} \tag{3.17a}$$

where k_x and k_ψ are arbitrary coefficients for the correction in calculation. Instead of (3.16a) we get the rearranged expression

$$y \equiv \alpha\Lambda^* = \cdots = \frac{\frac{B}{\sqrt{x}} - \frac{3}{2}\frac{x\psi}{A}}{\exp\frac{x^2\psi}{A}} \equiv F(x, \psi) \cdot e^{-\frac{x^2\psi}{A}}, \quad (x < 1) \tag{3.17b}$$

suitable for the quantitative determination of parameter from the exact numerical results (according to [39]).

We remark that Wolff [40] was the first who gave the theoretical treatment (by means of kinetic equation for distribution function) for finding $\alpha_{n,p}$'s: he calculated the dependence $\alpha(K)$ and he managed to get the analytical dependence of the type:

$$\alpha_{n,p} = A_{n,p} \cdot e^{-\left[\frac{b_{n,p}}{K(x)}\right]^m}, \quad (A \text{ and } b \neq f(K)), \tag{3.16b}$$

for silicon (with $m = 2$). There are various approaches both in [41] and [24] — 1961 to the relations similar to (3.16b) but with $m = 1$, while regarding low fields ($K \ll K'$) it is shown in [39] that $m = 1$ should be taken and by strong fields ($K \gtrsim K'$)—$m = 2$. We see that our rearranged expression (3.17b) has advantages because:

— it has an *adequate* physical interpretation for every parameter,

— it shows that the $\exp\left(-x^2\frac{\psi}{A}\right)$ *need not be a preponderant* factor for the change of α with K, but that may be $F(x, y) = F(K)$;

*) In [38] — 1965 the attention is paid also to the orientation of K in relation to the principal axis of the equi-energetic ellipsoid; the relation is roughly similar to (3.16a').

— the parameter A is *universal* (it is compared with [39] in [21] and $A=3$ is calculated); thus we have four (Λ^*, B, W^* and $\hbar\omega_0$ or ψ) parameters, which is more than in other works, with mostly three parameters, i.e. eqn. (3.17b) is *more correct*.

3° — *Multiplication factors* are best defined as the ratios of current at the output to the current of same carriers at the input of the transition (and not as the ratio of concentrations):

$$M_p=\frac{J_p(x_P)}{J_p(x_N)}\equiv\frac{J_{pN}}{J_{p0}},\quad M_n=\frac{J_{nM}}{J_{n0}};\quad \left(M_n'=\frac{n_M}{n_0},\quad M_p'=\frac{p_M}{p_0}\right), \tag{3.18a}$$

not only because $M_{p,N}$ is *not equal to* $M'_{p,n}$, but also because the *starting relation* (eqns. 2.2 without the recombination part and diffusion current)

$$-\frac{1}{e}\frac{dJ_n}{dx}=g_i+g_n'\equiv g_i+(n\nu_n+p\nu_p)=\frac{dJ_p}{dx}\cdot\frac{1}{e} \tag{3.18b}$$

is written in a simpler way by means of current*):

$$\frac{dJ_n}{dx}=\alpha_n J_n+\alpha_p J_p+eg_i=-\frac{dJ_p}{dx} \tag{3.18c}$$

with:

$$J_p=epv_{dp}\ \text{and}\ J_n=-env_{dn}$$

and it represents the linear differential equation:

$$\frac{dJ_n}{dx}-\alpha(x)\cdot J_n(x)=\alpha_p(x)\cdot J+eg_i,\quad \alpha\equiv\alpha_n-\alpha_p \tag{3.18d}$$

where, by *definition*:

$$\alpha_p=\frac{\nu_p}{|v_{dp}|}\ \text{and}\ \alpha_n=\frac{\nu_n}{|v_{dn}|},\quad (\alpha_{n,p}>0), \tag{3.18c}$$

under the condition that drift velocities ($v_{dp,n}$) *need not* be (for lower K) the saturation ones ($v_{sp,n}$). By solving (3.18d) and knowing that ([21]):

$$J=J_n(x)+J_p(x)=J_{nM}+J_{p0}=J_{n0}+J_{pM},\quad \frac{dJ_n}{dx}=-\frac{dJ_p}{dx},$$

$$(J=J_n+J_p=C^{te}); \tag{3.18f}$$

$$I_p(x,x_0)\equiv\int_{x_0}^{x}\alpha_p(y)\cdot\exp\left(-\int_{x_0}^{y}\alpha\,du\right)\cdot dy=I_n-1+\exp\left(-\int_{x_0}^{x}\alpha\,du\right),$$

according to the definition (3.18a) — left, we get finally:

$$M_n=\left(1+\frac{J_{p0}}{J_{n0}}I_{p\varphi}+G_{i\varphi}/J_{n0}\right)\Big/(1-I_{n\varphi}),$$

$$M_p=\left(1+\frac{J_{n0}}{J_{p0}}I_{n\varphi}\cdot e^{I\varphi}+G_{i\varphi}\,e^{I\varphi}/I_{p0}\right)\Big/(1-e^{I\varphi}\cdot I_{p\varphi}) \tag{3.18g}$$

*) As we have $v_{dn}>0$, v_{dp}, J_n and $J_p<0$ both for reverse voltage and for N-P transition as in Fig. 2.1b, reverse currents are considered here further on as *positive*: calculations are carried out with the *absolute values* of current.

ν_n and ν_p are the rate of ionizing collision acts *per carrier*.

where the integrals

$$I_{p\rho}=I_p(x_N, x_P);\quad I_{n\rho}=I_n(x_N, x_P);\quad I_\rho=I_y(x_N, x_P);$$

(3.18h)
$$I_n=\int_{x_P}^{x}\alpha_n\, e^{-I_y}dy,\qquad I_y=\int_{x_P}^{y}\alpha\, dn,$$

are *dependent* on the field, i. e. on the voltage applied.

Similar expressions are obtained at other places, but

— it is assumed in [42] that the generation does not change with x i.e. with K, so that instead of our

(3.18i)
$$G_{i\rho}=G_i(x_N, x_P),\quad G_i(x, x_P)\equiv e\int_{x_P}^{x} g_i(y)\, e^{-\int_{x_P}^{y}\alpha\, du}\cdot dy,$$

a simplified expression is obtained, which need not always be correct; Čuenkov considers that $g_i=0$ (in [31]) and he neglects J_{p0};

— M_n and M_p can be immediately seen from (3.18g) to be *different* not only in the general case but even when $\alpha_n-\alpha_p=\alpha=0$: ionization by electrons and holes being *the same*, ($M_p=M_n$ only for $J_{p0}=J_{n0}$);

— if we assume that approximately

(3.19a)
$$\alpha_p=k_\alpha\alpha_n,\quad k_\alpha\neq f(K)$$

which does not result from (3.17b), but it might be accepted for a ,,narrower" region of K's, and we obtain

(3.19b)
$$M_n=\frac{k_\alpha-1}{k_\alpha-e^{-I\rho}}$$

(as for instance in [25] — 1956) *only* if J_{p0} is neglected together with g_i.

c. — Breakdown criterion

1° — In all the papers known to us, the breakdown is considered to take place when the corresponding value, e.g. current, becomes extremely high at the ,,excitation" (for instance voltage or the field) reaching the corresponding breakdown (critical) value:

(3.20a)
$$I,\ J\to\infty\ \text{ for }\ U=U_b\ \text{ or }\ K=K_b.$$

Nevertheless, if we consider it physically, at least two objections can be made:

— the ,,excitation" in one process *approaches* a certain value, hence $U\to U_b$ or $(U_b-U)\to+0$,

— and apart from that, *new* effects may originate from the tendency of the sudden increase of I, and they may be of great importance in that situation, so that the *further increase* is slowed down to a certain extent.

Hence, the process $I\to\infty$ may not even take place, but a situation when the slightest *change* of excitation causes an extremely great change of I may take place in the common effect of the ,,old" and ,,new" processes:

(3.20b)
$$\frac{dI}{dU}\to\infty\ \text{ when }\ U\to U_b,$$

which is physically more adequate than (3.20a); when $U \to U_b$ the current may have the *finite value* $I = I_b$.

2° — By applying the above to the avalanche breakdown, the breakdown condition is obtained from (3.18g):

$$1 - I_{n\rho} = 0 \quad \text{or} \quad 1 - e^{I\rho} \cdot I_{p\rho} = 0, \tag{3.20c}$$

with both the conditions identical, which can be easily seen when using (3.18f) — below. Thus, the *breakdown criterion* is finally:

$$\int_{x_{Pb}}^{x_{Nb}} \alpha_n \, e^{-\int_{x_{Pb}}^{y} \alpha \, du} \, dy = 1 = \int_{x_{Pb}}^{x_{Nb}} \alpha_p \, e^{\int_{y}^{x_{Nb}} \alpha \, du} \cdot dy; \quad \int_{x_{Pb}}^{x_{Nb}} \alpha_n \, dx = \frac{\ln k_\alpha}{k_\alpha - 1}, \tag{3.20d}$$

where the final expression is valid *only for* $k_\alpha = const$.

By applying the above to the calculation of the minimum width of SC ($\Lambda_{\rho min}$) needed for the start of the avalanche breakdown (but not that of Zener), according to (3.20d) and (3.17b) we have got the quantitative relation $\Lambda_{\rho b}(K_b)$ in [21] and calculated $\Lambda_{\rho \min} \approx 500$ Å for Si and Ge, which *agrees well* to the well-known fact that Zener's breakdown takes place at small widths of about 400 Å ([25] — 1960).

d. — Breakdown regime current

1° — Let us assume that $k_\alpha = \alpha_p/\alpha_n = C^{te}$ is valid. Then (3.18f), (3.18g) and other expressions become somewhat simpler, and the following expression is obtained after a detailed analysis (in [21]):

$$J = \frac{k_\alpha - 1}{k_\alpha} \cdot \frac{\dfrac{D_{nP}}{L_{nP} P} + \dfrac{D_{pN}}{L_{pN} N} k_\alpha \cdot \mathcal{B}(U)\,(1 + \gamma_{gd})}{1 - \mathcal{B}(U)} \cdot e n_i^2 \tag{3.21a}$$

where, by applying (3.17b), $\mathcal{B}(U)$ is:

$$\mathcal{B}(U) = \frac{1}{k_\alpha} \exp\left[(k_\alpha - 1) \int_{x_P}^{x_N} \alpha_n \, dx \right] = \cdots =$$

(3.21b)

$$= {}^2\mathcal{B}(U) = \frac{(U_b/U)^{\frac{M_2}{2}(k_\alpha - 1)}}{\xi_b^{2M_2(k_\alpha - 1)/3}} \cdot \frac{\exp\left[\left(\dfrac{U}{U_b}\right)^{3/4} \cdot (k_\alpha - 1) \cdot \dfrac{2M_2}{3} \xi_b \right]}{k_\alpha \cdot \exp\left[\dfrac{2M_2}{3} (k_\alpha - 1) \right]} \equiv \frac{C_b}{k_\alpha} U^{-b_1} e^{b_2 U^{3/4}},$$

and where ${}^2\mathcal{B}(U)$ corresponds to the *abrupt asymmetrical transition* with the constants:

$$M_2 = \frac{3 \varepsilon \varepsilon_0 W^* \psi_n}{2 e^2 (\Lambda^*)^2 A}, \quad \xi_b = \left(\frac{K_b}{K_1}\right)^{3/2}, \quad K_1 = K_n \left(\frac{3 \psi_n}{2 A B_n}\right)^{2/3}; \tag{3.21c}$$

K_b and U_b are the *breakdown field and voltage* which are obtained from (3.20d):

$$\frac{N}{N_0}=\frac{k_\alpha-1}{\ln k_\alpha}\left\{[(\xi_b-1)-\ln\xi_b]+\sum_p\right\},\quad K_b^2=\frac{2eNU_b}{\varepsilon\varepsilon_0} \tag{3.21d}$$

with

$$N_0=\frac{2}{3}M_2N=\frac{\varepsilon\varepsilon_0 W^*\psi_n}{e^2(\Lambda^*)^2A},\quad a=\frac{K_0}{K_1}, \tag{3.21e}$$

$$\sum_p\equiv\sum_{p=1}^{\infty}\frac{3a^{2p}(-1)^p}{4\cdot p!}\left(\frac{1-\xi_b^{1-\frac{4}{3}p}}{p-3/4}-\frac{1-\xi_b^{-\frac{4}{3}p}}{p}\right).$$

The ratio γ_{gd} is

$$\gamma_{gd}=\frac{1-e^{l\rho}}{k_\alpha-1}\cdot\frac{L_{pN}N}{\bar{\alpha}_n\cdot n_i\tau_\infty D_{pN}}\approx\frac{N}{n_i\sqrt{\tau_\infty D_{pN}\cdot\bar{\alpha}_n}},\quad \bar{\alpha}_p=k_\alpha\bar{\alpha}_n. \tag{3.21f}$$

We see that the current is expressed through our *basic* parameters (W^*, Λ^*, B and ψ), *technological* parameters (N, τ, ...) and intrinsic properties (ε, n_i, ...).

2° — Let us analyse our expression (3.21a) now, and compare it with the results of the others.

By assuming that $\alpha_{n,p}$ is given (empirically) with the expression $\alpha_{n,p}=C_{n,p}K^r$ Vul has obtained the expression:

$${}^{o}\mathcal{B}(U)=\frac{1}{k_\alpha}(k_\alpha)^{\rho},\quad \rho=\left(\frac{U}{U_b}\right)^{\frac{r+1}{2}} \tag{3.22a}$$

for the *abrupt transition*. We have put the above index „o", because $\alpha\sim K^r$ is really a zero approximation.

Starting from the approximation $\alpha_n\sim e^{-\frac{b_n}{K}}$, which might be somewhat more correct but physically it cannot be justified either (see point **b.**), and after neglecting some facts not quite in an adequate way, Armstrong obtained [42]

$${}^{1}\mathcal{B}(U)=\left(\frac{K_{max}}{K_b}\right)^m\cdot\frac{\exp(b_n/K_b)}{\exp(b_n/K_{max})},\quad K_{max}\equiv|K|_{max}\equiv K_m \tag{3.22b}$$

with $m=0$, $1/2$ and 1 for **P-I-N**, linear and abrupt transition, respectively. Both according to (3.22a) and according to (3.22b) the constants $C_{n,p}$, r and b_n cannot be adequately linked to the real parameters of the avalanche ionization ($\hbar\omega_0$, W^*, Λ^* and similar).

Our expression (3.17b) for $\alpha\Lambda^*$ is already partially checked*) (for strong fields); the value $x^2\psi/A$ ($\psi\approx0.02$, $A=3$, $x<10$, $K\gtrsim K_0$) can be easily seen then

*) We have got (3.21d) on the basis of (3.17b). Then $U_b=26.6$ V is obtained for Ge for $N=10^{16}$ cm^{-3}, while according to the experimental data $U_b=27$ V! We have further developed the idea of getting dependence $U_b(N)$ in [46] for Ge, Si (and GaAs) and obtained satisfactory agreement.

not to exceed the value of about 0.7, so that it is possible to write approximately (for $e^z \approx 1+z$)

$$\Lambda_n^* \alpha_n = \frac{B_n}{\sqrt{K_n}} \cdot \sqrt{K} - \frac{3\psi_n K_n}{2A} \cdot \frac{1}{K} - \frac{B_n \psi_n}{A} K_n^{3/2} \cdot \frac{1}{K^{3/2}}, \tag{3.22c}$$

where we introduced separate constants B_n, ψ_n, K_n and Λ_n^* for electrons, because B depends on p_0 i.e. on w_0 from (3.15c) i.e. on the sorts of carriers and materials, and we have left A to be „universal" ($A=3$). In that way we have obtained (3.21b) for ${}^2\mathcal{B}(U)$ where all the parameters are physically clear; b_1 and b_2 are either both positive (for Ge) or both negative (for Si). By comparing ${}^2\mathcal{B}(U)$ with ${}^1\mathcal{B}(U)$ according to (3.22b) we see that:

— $\mathcal{B}$'s are rapidly varying functions of U near U_b in both the cases,

— both the expressions contain the product in the form $U^{-b_1} e^{f_2(U)}$,

— but b_1's *differ* in magnitude and they may differ in sign, while the functions $f_2(U)$ can differ very much.

The assumption $\alpha_n = \alpha_p$ i.e. $k_\alpha = 1$ (used for instance in [43]) is very rough, but if it is accepted, the result obtained is very interesting:

$$M_n = \frac{1 + \frac{J_{p0}}{J_{n0}} \cdot I_{n\varrho_0} + \frac{G i_\varrho}{J_{n0}}}{1 - I_{n\varrho_0}}, \quad I_{n\varrho 0} \equiv \int_{x_P}^{x_N} \alpha_n \, dx, \tag{3.22d}$$

and it is possible to see from (3.21b) and from (3.22d) that the following expression is *not obtained* for M_n:

$$M_{n,p} \sim \frac{1}{1 - \left(\frac{U}{U_b}\right)^{n_{n,p}}} \tag{3.22e}$$

which has been used empirically by many authors. The expression (3.22e) cannot be obtained from (3.22b) either, but only from (3.22a) and that exclusively for $\alpha_n = \alpha_p$.

That the coefficients $n_{n,p}$ in (3.22e) are not constants is proved even earlier in [17] — *Vol.* **1,** pp 294 to 310 (Fig. 10).

For *smaller reverse* voltages, $\mathcal{B}(U)$ is approximately $1/k_\alpha$ and γ_{gd} according to (3.21f)

$$\gamma_{gd} = \cdots = \frac{0}{0} = \frac{L_{pN} \cdot N}{n_i \tau_\infty D_{pN}} (x_N - x_P) \tag{3.22f}$$

and then (3.21a) transforms, roughly, in an expression obtained in [22] and [32] for reverse polarization neglecting the avalanche effect, where ${}^\circ J$ depends upon U, only through $(x_N - x_P) = \Lambda_\varrho$.

By developing $\mathcal{B}k_\alpha$ into the series ($\mathcal{B}k_\alpha \approx 1 - I_\varrho + \cdots$) in [21], the more correct expression for current (${}^1J$) is obtained. It enables the *conclusion* that the influence of multiplication is negligeable for $U \leqslant 0.1\, U_b$, but for $U = 0.6$ up to 0.8 of U_b this influence is *considerable*, and that happens far from the actual breakdown.

Chapter 4. — VERIFICATION OF RESULTS AND FURTHER ANALYSIS

The accuracy of the expressions derived: for the distribution of potential $\varphi(x)$, the temperature dependence of the width of space charge, the ionization coefficient, the „recombination" function I_3^* and the expression (3.11 b) for $G(y)$, will be estimated.

Various methods, either those based on our experiments or experiments from other authors or those based on our calculation or on data of other authors, will be used.

In the further analysis we shall try to find out, whenever it is possible, either the applicability of our results or the consequences.

4.1. — Estimation of the accuracy of the potential relation and its application

In Section 2.2 point *c* the expressions for potential function $\tilde{\varphi}(x)$ were derived and comments have been given, respectively. Now it is necessary to make a qualitative or semi-qualitative estimation of the accuracy of these results.

At the beginning the *errors* in those expressions similar to (2.21 e) and (2.22) will be estimated and our expressions $\tilde{\varphi}(x)$ will be *compared* with the exact ones in relation to asymmetric abrupt transition ($P>N$).

The relation obtained in [23] has been taken for comparison, because a precise calculation with a calculating machine was done.

4.1.1. — THE ESTIMATION OF ERROR

a. — General

The mean-value theorem (for integrals), will be used for the estimation of the error, using the way like in [23] and the same symbols like in Section 2.2.

For an abrupt transition $\mathcal{N}_u = \mathcal{N}$ is valid and $Z_u^2 = 0$, so that the denominator of the integral (2.21 c) can be written as

$$[\cdots]^{1/2} = -\tilde{\alpha}\,(\cosh\tilde{\varphi}_{u/2})^{1/2}\left(1 - \frac{\tanh\tilde{\varphi}_u}{3}\tilde{\alpha} + \frac{\tilde{\alpha}^2}{12} - \frac{\tanh\tilde{\varphi}_u}{60}\tilde{\alpha}^3 + \cdots\right)$$

$$= -\tilde{\alpha}\left(1 - \frac{\tanh\tilde{\varphi}_u}{3}\tilde{\alpha} + \frac{\tilde{\alpha}^2}{12}\right)^{1/2} \cdot [1 + F(\tilde{\alpha})]^{1/2} \cdot \left(\frac{\cosh\tilde{\varphi}_u}{2}\right)^{1/2},$$

where F is a „residual" functional ($|F|$ is smaller, or much smaller than one), and the integral can be written as the definite one

$$\pm\frac{x-x_c}{\Lambda'_{Di}} = \int_{\tilde{\alpha}_c}^{\tilde{\alpha}} \frac{f(\tilde{\alpha})\,d\tilde{\alpha}}{\sqrt{1+F}} = \frac{1}{\sqrt{1+\bar{F}}}\int_{\tilde{\alpha}_c}^{\tilde{\alpha}} f(\tilde{\alpha})\,d\tilde{\alpha} \equiv \frac{\Phi(\tilde{\alpha})}{\sqrt{1+\bar{F}}}, \quad \text{with} \tag{4.1 a}$$

$$f(\tilde{\alpha}) \doteq \tilde{\alpha}^{-1}\cdot\left(1 - \frac{\tanh\tilde{\varphi}_u}{3}\tilde{\alpha} + \frac{\tilde{\alpha}^2}{12}\right)^{-1/2}, \quad \frac{d\Phi}{d\tilde{\alpha}} = f(\tilde{\alpha}); \quad \alpha = \varphi_u - \varphi, \tag{4.1 b}$$

where φ_u is the „asymptotic" value of potential for P-or N-region, $\alpha_c = \varphi_u - \varphi_c$; $\bar{F} = F(\tilde{\alpha}_1)$, $\alpha_1 = C^{te}$ and its value is between α_c and α. Our expression (2.22) is approximate (it was assumed that $\bar{F} = 0$). Let us assume that (2.22) is exact for $\Phi(\tilde{\beta}) = \pm \frac{x - x_c}{\Lambda'_{Di}}$ so that*) it is:

$$\alpha = \beta + \delta\alpha, \quad \Phi(\tilde{\alpha}) = \Phi(\tilde{\beta} + \delta\tilde{\alpha}) = \Phi(\tilde{\beta}) + \delta\Phi = \sqrt{1 + \bar{F}}\,\Phi(\tilde{\beta}) \approx \left(1 + \frac{\bar{F}}{2}\right)\Phi(\tilde{\beta})$$

| (4.1 c) |

$$\therefore \delta\Phi = \frac{\bar{F}}{z}\Phi(\tilde{\beta}) \approx \frac{d\Phi}{d\tilde{\alpha}}\delta\tilde{\alpha} \quad \text{or} \quad |\delta\tilde{\alpha}| \approx \left|\frac{\bar{F}}{2}\right| \cdot \frac{\Phi(\tilde{\alpha})}{|f(\tilde{\alpha})|},$$

and it is necessary that the error is smaller than $|\delta\tilde{\alpha}|$ for $\tilde{\alpha}$ in the interval $(\tilde{\alpha}_c, 0)$.

This general expression *is not derived* in [23], and from (4.1 c) one can find that it is $\alpha \ln \frac{\alpha_c}{\alpha}$ for $|\Phi|/|f|$ if expanding $\cosh\tilde{\varphi}$ one stops at square term (then it is $f(\tilde{\alpha}) = 1/\tilde{\alpha}$).

The estimation for the largest values for $|\bar{F}|$ and $|\Phi|/|f|$ was made in [21] so that taking this into account and expression (4.1 c) *the error* and *the biggest* value for $|\tilde{\alpha}_c|$ are:

| (4.1 d) |

$$|\delta\tilde{\alpha}| \leqslant \frac{|\tanh\tilde{\varphi}_u|}{80 \cdot e}\,|\tilde{\alpha}_c|^4 \quad \therefore \quad |\tilde{\alpha}_c| \leqslant \left(\frac{|\delta\tilde{\alpha}| \cdot e \cdot 80}{|\tanh\tilde{\varphi}_u|}\right)^{1/4}$$

Using our notation for expression in [23] one can get that it is $|\delta\tilde{\alpha}| \leqslant \frac{|\tanh\tilde{\varphi}_u|}{6e}|\tilde{\alpha}_c|^2$, which means that *for the same* interval $|\tilde{\alpha}_c|$ the relations (2.22) will have a $(3|\tilde{\alpha}_c|^2/40)$ times *smaller* error from the respective „exponential" relationship $\left(\tilde{\alpha} \sim \exp\frac{-x}{\Lambda'_{Di}}\right)$. That means that our expression will be *more accurate* up to the value $|\tilde{\alpha}_c| = \sqrt{40/3} = 3.65$ i.e. that *the interval* $|\tilde{\alpha}_c|$ using our expression will be *larger* until the tolerated error is fulfilling the following condition:

(4.1 e) $$|\delta\tilde{\alpha}| < \frac{20}{9e}|\tanh\tilde{\varphi}_u|.$$

One can *conclude* from the above that our relations are practically *always more accurate* and that the interval, in comparison to the expression given in [23], is *extended* for

(4.1 f) $$|\tilde{\alpha}_c|_{(2.22)} - |\tilde{\alpha}_c|_{[23]} = \left(\frac{|\delta\tilde{\alpha}|\,80\,e}{|\tanh\tilde{\varphi}_u|}\right)^{1/4} - \left(\frac{|\delta\tilde{\alpha}|\,6\,e}{|\tanh\tilde{\varphi}_u|}\right)^{1/2};$$

for instance, if $\left|\frac{\delta\tilde{\alpha}}{\tanh\tilde{\varphi}}\right| = 10\%$, it is $(2{,}17 - 1{,}27) \approx 1$.

We are going to use those results on some abrupt transitions.

*) Minus sign is related to the N-region ($x \geqslant 0$) and plus to P-region ($x \leqslant 0$); the same is valid in (4.1 a).

b. — P-I Transition

We are not going to discuss the distribution of potential in I-region ($x>0$), because it can be expressed with an elementary function (in equation 2.21c, $Z_u^2=0$, $\tilde{\varphi}_u=\tilde{\varphi}_I=0$ and $\mathcal{A}=\mathcal{A}_I=0$).

For P-region of an P-I transition $|\tilde{\alpha}_c|$ is *smaller* than $(\tilde{\varphi}^\circ-\tilde{\varphi}_P)$ so that from TB.d for $N=0=\tilde{\varphi}_N$ one can get easily

$$\tilde{\varphi}^\circ=\frac{2n_i(\cosh\tilde{\varphi}_P-1)+\varphi_P P}{P}\quad\therefore\ |\tilde{\alpha}_c|\leqslant|\tilde{\alpha}^\circ|=\tilde{\varphi}^\circ-\tilde{\varphi}_P=\frac{\cosh\tilde{\varphi}_P-1}{\sinh(-\tilde{\varphi}_P)} \tag{4.2a}$$

and the *largest* error in *the whole* P-region (that means on the edge of the transition) is equal, using equation (4.1 d):

$$|\delta\tilde{\alpha}|\leqslant\frac{|\tanh\tilde{\varphi}_P|}{80\,e}|\tilde{\alpha}^\circ|^4=\frac{(\cos\tilde{\varphi}_P-1)}{80\,e\cosh\tilde{\varphi}_P|\sinh\tilde{\varphi}_P|^3}\equiv y(\tilde{\varphi}_P)\leqslant\frac{1}{80\,e}=0{,}46\%. \tag{4.2b}$$

One can show that the function $y(\tilde{\varphi}_P)$ has no extremes and that it grows monotonously, approaching (assymptotically) the above value, when P/n_i and $|\tilde{\varphi}_P|$ is approaching infinity — i.e. descriptively: for strongly doped P-region. Using [23] one can see that the error would be about 6,1%.

We should remark that the analysis leading towards the relation of type (3.2b), *has not been done* in [23] and as far as we know nowhere else.

c. — Symmetrical abrupt P-N transition

Now $\tilde{\varphi}_N=-\tilde{\varphi}_P=\sinh^{-1}\frac{N}{2n_i}$ and the biggest value for $|\tilde{\alpha}_c|$ is $\tilde{\varphi}_N$ so that for the *whole* transition the largest error will be like in T4.1, depending on the doping level N/n_i. This fact, once again confirms what was said in section **a**:

$|\delta\tilde{\alpha}|=f\left(\frac{N}{n_i}\right)$ *for the whole symmetrical transition* T 4.1

N/n_i	$\frac{5.78}{100}$	0.1	0.26	1	5.78	10	38.45*)	57.8
$\varphi_N=\sinh^{-1}\frac{N}{2n_i}$	$\frac{2.89}{100}$	0.05	0.13	0.483	1.782	2.31	3.65	4.055
100 tanh φ_N	2.89	5	12.9	44.9	94.6	99	100	100
$\|\delta\alpha\|$, eqn. (4.1d) [%]	$\approx10^{-8}$	$1.5\cdot10^{-7}$	$1.7\cdot10^{-5}$	0.011	4.4	13	81	124
ref. [23]	$1.5\ 10^{-4}$	$7.7\cdot10^{-4}$	$1.3\cdot10^{-2}$	0.64	18.5	32.4	81	100

*) For this value the errors found, using [23] and (4.1 d), are equal. The other values different from decade units are taken for comparison from [23].

— our approximation is far better than the „exponential" one for „reasonable" values of errors;

— for higher doping levels (higher than 10), even our approximation cannot cover $\tilde{\varphi}$'s for the whole region of transition — in the neighbourhood of $x=0$ there exists another distribution of potential.

d. — Asymmetrical P-N transition

Let us assume that the *degree of asymmetry* $\gamma=\frac{P}{N}=\frac{\sinh|\tilde{\varphi}_P|}{\sinh|\tilde{\varphi}_N|}$ $(P>N)$, is much bigger than 1, because if not, the situation is very similar to symmetrical transition which already has been analyzed. Besides that, let us assume that the doping level ot N-region is much higher than 1 $\left(\frac{P}{n_i}>\frac{N}{n_i}\gg 1\right)$, because in the opposite situation it would behave almost like I-space (see section **b**). The general expression for $|\tilde{\alpha}^\circ|$ is from TB.d:

$$|\tilde{\alpha}^\circ|=\tilde{\varphi}^\circ-\tilde{\varphi}_P=\frac{\cosh\tilde{\varphi}_P-\cosh\tilde{\varphi}_N+\sinh\tilde{\varphi}_N(\tilde{\varphi}_N-\tilde{\varphi}_P)}{\sinh\tilde{\varphi}_N-\sinh\tilde{\varphi}_P}\equiv z(\tilde{\varphi}_N,\tilde{\varphi}_P). \tag{4.3 a}$$

1° — In the *P-region* it would be useful to estimate the *largest value* for $|\tilde{\alpha}^\circ|$. It was shown in [21] that this function has *no extremes* in the proper sense, which is physically clear, but for $\tilde{\varphi}_N$ the condition $\frac{\partial z}{\partial\varphi_P}=0$ *could be* fulfilled in interval between 1 and 1.91 which means that there $z(|\tilde{\varphi}_P|)$ is not a monotonic function. Meanwhile, for bigger γ's, $|\alpha^\circ|\equiv z$ is *tending* towards 1 *not depending**) on the doping level in N-region (only $\tilde{\varphi}_N\ll-\tilde{\varphi}_P$ must hold), which one can see easily from expression (4.3 a).

From all this, one can *conclude* that

— for $\tilde{\varphi}_N\leqslant 1.91$ i.e. for $\frac{N}{n_i}<6.6$ for *any* degree of asymmetry, $|\tilde{\alpha}^\circ|$ is not larger than 1.91, so that by expression (4.1 d) is

$$|\delta\tilde{\alpha}|\leqslant\frac{|\tanh\tilde{\varphi}_P|}{80\,e}(1.91)^4<\frac{(1.91)^4}{80\,e}\approx 6\%; \tag{4.3 b}$$

in the *whole* P-region (using [23] the error would be even 22%);

— the largest error, for the whole P-region and any γ's, the doping levels being higher than 6.6 ($\tilde{\varphi}_N>1.91$), would be

$$|\delta\tilde{\alpha}|<\frac{|\tanh\tilde{\varphi}_P|}{80\,e}|\tilde{\varphi}_N|^4, \tag{4.3 c}$$

while for a strong asymmetry ($\gamma\gg 1$) it would be only:

$$|\delta\tilde{\alpha}|<\frac{1^4}{80\,e}\approx 0{,}5\%. \tag{4.3 c'}$$

*) We can note that, as far as we know, nobody has mentioned this conclusion before.

To compare this with the results from [23] let us assume that $|\delta\tilde{\alpha}| = 10\%$ and then $\tilde{\alpha}^{\circ} = 2{,}16 \approx 2$. That means that for $\gamma = 1$ one can go up until the doping level reaches $\frac{N}{n_i} \approx 10$ (or for $\frac{N}{n_i} \approx 10$ with *any* γ). But using the assumption made at the beginning of this section, one can suppose that approximately

(4.3e) $$2\sinh\tilde{\varphi}_N \approx e^{\tilde{\varphi}_N} \approx 2\cosh\tilde{\varphi}_N \quad \text{and} \quad 2\cosh|\tilde{\varphi}_P| \approx e^{|\tilde{\varphi}_P|} \approx 2\sinh|\varphi_P|,$$

especially because we are interested now in region $|\tilde{\varphi}_P| > \tilde{\varphi}_N > 2$. In [21] it was shown that between the *smallest* degree of asymmetry γ_m (for the error of 10%) and doping of N-region $\frac{N}{n_i} \approx e^{\tilde{\varphi}_N}$ there is an approximate relation which could be written like:

(4.3f) $$\frac{N}{n_i} \approx \frac{\exp\frac{\gamma_m + 3}{2}}{\sqrt{\gamma_m}}, \quad \text{or} \quad \gamma_m = \ln\gamma_m + 2\tilde{\varphi}_N - 3.$$

One can see that γ_m *depends* on the doping level too, and *increases* with $\tilde{\varphi}_N$, so that the statement made in [23], which says that the exponential approximation is valid for the whole P-region when $\frac{P}{N} > 10$ (with 10% error) is *not true* even for our more exact approximation. From (4.3f) it follows that the above is correct *only* for $2\varphi_N = 10 - 2.3 + 3 = 10.7$ or for $\frac{N}{n_i} \leqslant 200$.

2° — In *N-region* our approximation need not be valid near the transition (about $x \approx 0$, range II in Figure 2.3b), as it was said in point **c.** — Section 2.2. So the space beginning with $x > x_i$ — regions III and IV in Figure 2.3b is interesting here and, also the question of the magnitude of the error $|\delta\tilde{\alpha}|$ for $x \approx x_i$ or $x \leqslant x_i$ (when $\tilde{\varphi}_c \approx 0$ or $\tilde{\varphi}_c < 0$). Now $\tilde{\alpha}_c \equiv \tilde{\varphi}_N - \tilde{\varphi}_c \geqslant \tilde{\varphi}_N$ and $\tilde{\varphi}_u = \tilde{\varphi}_N$; using this and equation (4.1d) one can easily make the analysis. *The numerical values* are the same as in T 4.1 but they are *not valid* for the whole N-region but only up to $x \approx x_i$ i.e. for $0 \leqslant \tilde{\varphi} \leqslant \tilde{\varphi}_N$. It is even better *not to connect* the quantity N/n_i with the largest $\tilde{\alpha}_c$ for the given error.

e. — **The conclusion** could be that our approximation (2.22) *always* gives *better results* than the usual „exponential" one, covering practically the whole P-region of asymmetrical P-N transitions, with the error less than 10% for $\gamma > 10$ and $N/n_i \leqslant 200$, and it is *almost exact* (0.46%) for the same region of P-I transition. Roughly, $|\delta\tilde{\alpha}|$ is smaller than 15% for $|\tilde{\alpha}_c| \leqslant 2 \div 3$.

Taking into account the approximative relations

(4.4a) $$\frac{x}{\Lambda'_{Di}} = C_1 + \ln|\tilde{\alpha}| \quad \text{and} \quad |\tilde{\alpha}| = C_2 + \ln p \quad \text{or} \quad |\tilde{\alpha}| = C_3 + \ln n,$$

where n and p are the respective concentrations of carriers and x is the distance, one can see that

(4.4b) $$\frac{|\delta x|}{\Lambda'_{Di}} = \frac{|\delta\alpha|}{\alpha} \quad \text{and} \quad \frac{|\delta p|}{p} = |\delta\tilde{\alpha}| = \frac{|\delta n|}{n},$$

so that $|\delta x|$ can be even *smaller* than $|\delta\tilde{\alpha}|$, while $|\delta\tilde{\alpha}|$ is the *relative error* for carrier concentrations.

4.1.2. Calculation of variation $\tilde{\varphi}(x)$ for some cases; its application to x_i

a. — General

In [23] the *exact variations* of $\tilde{\varphi}(x)$ were given by using the data obtained with electronic calculating machine („STRELA"), for several cases (A, B and C as given in T 2.1).

b. — The treatment

— At the beginning we calculated the relation $\tilde{\varphi}=\tilde{\varphi}_{IV}(x')$, using the data given in T 2.1 for all three cases and our expression (2.22) (the potentials have indexes IV). The constant C_2' was calculated temporarily from the condition $\tilde{\varphi}_{IV}=0$ for $x'=0$, and then matching point $\tilde{\varphi}_{ma}$ was chosen in such a way that $\varphi_{IV}(x')$ is identical with the exact value $\tilde{\varphi}(x)$ which was calculated in [23] and then the „translation" was done to the „real" values of x:

$$x=x-a\Lambda_{Di} \tag{4.5a}$$

and the matching for cases A, B and C was done for

$$\tilde{\varphi}_{ma}=-0.87;\ 1.3 \text{ and } 8.61 \text{ so that } a=2.45;\ 0.75 \text{ and } 3.28\cdot 10^{-4} \tag{4.5b}$$

respectively.

— For region II the approximations from TB.b were used where the potentials were marked with indexes II and IIe, respectively.

— Besides that, we wanted to examine the exactness of the approximation of integral in (2.21c) for *region* II with an elliptic integral using TB.c (index „*el*").

All the variations of $\tilde{\varphi}(x)$ for the region II *start* from the point $\varphi^\circ\equiv\varphi(0)$, because this potential is exactly known (see TB.d and T 2.1).

For P-region the relations $\varphi(x)$ *were not calculated*, because we have shown already (Section 4.1.1) that our approximation was quite exact*).

c. — Results and conclusions

The relations $\tilde{\varphi}(x)$ for the cases A, B and C were calculated in [21] for more than 20 points (the calculations were done by M. Smiljanić — Institute of Physics, Beograd).

We are not giving the numerical data here because they are too extensive, so we have shown them as diagrams in Figures 4.1 a, b and c. The symbols are the same as in points **b** and **c**.

— full line („computer") — exact data obtained from [23];

— Sh-so called Shockley approximation or the „space charge" approximation- $\varphi(x)$ is then a parabola (see Sec. 2.2c).

The analysis of the dependences in Figure 4.1 gives the following *conclusions:*

---

*) One can see from T 2.1 that $|\tilde{\alpha}_c|=|\tilde{\alpha}^\circ|\equiv\tilde{\varphi}^\circ-\tilde{\varphi}_P$ is equal to unity for all three cases, which confirms our statement formulated in sec. 4.1.1 (that $\tilde{\alpha}^\circ\to 1$ for $\gamma\gg 1$) once again.

1° — Our expression (2.22) can approximate the variation of $\tilde{\varphi}(x)$ in an excellent way in almost the *whole* region for a *small doping* level of N-region Figure (4.1) — for a value of x which is much smaller than x_i that means even *out of the intermediate region*, and not depending on the fact that the degree of asymmetry $\gamma = \frac{P}{N}$ is big. This is in *quantitative agreement* with the analysis of errors made in Section 4.1.1., where one can see from (4.1d) that it is $|\delta\tilde{\alpha}| = |\delta\tilde{\varphi}| \sim \tanh \tilde{\varphi}_N$.

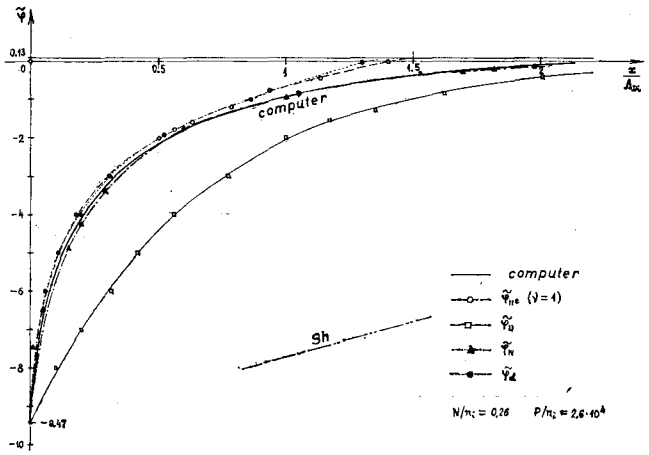

Fig. 4.1a. — Dependences $\tilde{\varphi}(x)$ for the case A;

The rest of our approximations: $\tilde{\varphi}_e$ and $\tilde{\varphi}_{el}$, *except* $\tilde{\varphi}_{II}$, are *quite satisfactory* in their regions — for $0 \leqslant x < \frac{x_i}{2}$. Disagreement in the variation of $\tilde{\varphi}_{II}$ can be explained as a consequence of choosing incorrect values for $\tilde{\varphi}_1$ (for all cases we have taken $\tilde{\varphi}_1 = -\tilde{\varphi}_N$).

2° — For the „*medium*" doping levels (Fig. 4.1b) one can use the conclusion from 1°, but with a bit *weaker* note.

3° — The results from Fig. 4.1c, which represents the *high* doping level of N-region, show that $\tilde{\varphi}_{IV}$ is still a good approximation for its region — for $\tilde{\alpha}_c \equiv \tilde{\varphi}_N - \tilde{\varphi}_c$ about 2—3, and that $\tilde{\varphi}_{II}$ is the best dependence in region II, (now the choice $\nu = 1$ is almost not important); here one should take a special approximation for region III.

4° — *All our relations* for $\tilde{\varphi}_{II}$, $\tilde{\varphi}_{IV}$, $\tilde{\varphi}_{IIe}$ and $\tilde{\varphi}_{el}$ give *far better* approximations than the „parabolic" one (Sh — Fig. 3.1) and they are *nearer* to the exact variation, even taking into account the ratio $\frac{N}{n_i}$ and a big disagreement for the case C. The advantage is that they have *analytical* character which was emphasized previously (in section 2.2b).

5° — Our relations could be improved, for instance with a better choice of parameters (φ_1, ν, φ_c etc.), better choice of a „residual" function $F(\alpha)$ in

(4.1a) and similar. This problem will not be analysed further, here; it could be the subject for a separate study.

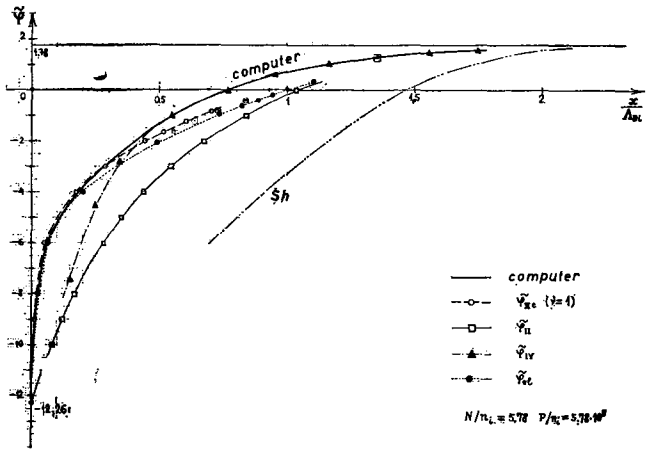

Fig. 4.1b. — Dependences $\tilde{\varphi}(x)$ for the case B;

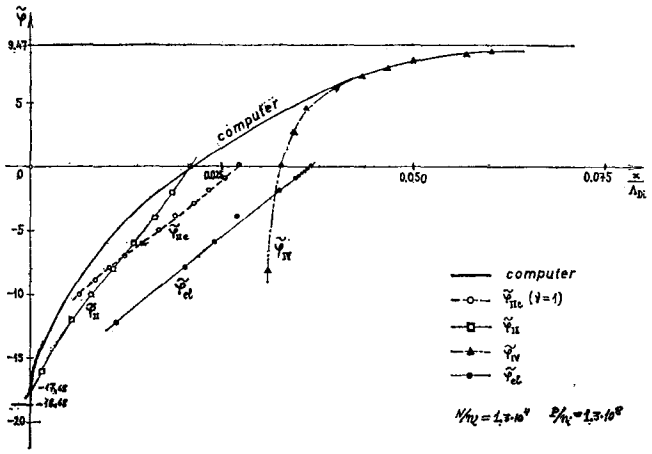

Fig. 4.1c. — Dependences $\tilde{\varphi}(x)$ for the case C.

6° — *The application on calculating the point of inversion* x_i given in Section 2.2c leads towards the results*) given in T 2.1. The last line in T 2.1 (x_i) relates to the *exact* values for x_i; the values which are the *nearest* to those are denoted by baldface fugures. On this basis we made a gradation of approximations for x_i in T 4.2. For the *best* agreement the error is not larger than about 15% for all cases. The „Elliptic" approximation („*el*") and Shockley ($^{\circ}x_i$) ones are not included in T 4.2 because they give the biggest disagreement (even bigger —45% and 224%).

*) Detailed calculation has been done in the Institute of Physics, Beograd, by M. Smiljanić.

The order of relations for x_i, by accuracy — T 4.2

Case	The best	A bit worse	Worse	Bad
A	(IV_e)	(IV)	(II)	(II_e)
B	(IV_e)	(IV)	(II_e)	(II)
C	(II) and (IV)	(IV)	(II_e)	(IV_e)

Using the above one can make the following *conclusion:*

— For „A" and „B" cases a better approximation will be that with „connection", because $\tilde{\varphi}_c < 0$, and that is *quite well confirmed* by the results from T 2.1 and T 4.2. The fact that (IVe) is better than (IV) one can explain with a too free choice of parameters φ_1 ($\varphi_1 = -\varphi_N$) in TB. (see point 1° too).

— For case „A" where there is a small doping level of N-region and even B, *relatively big* departures happen, which are obtained using the best expression TB.e, since the relation for φ_c gives too large absolute values, (for A it is $\tilde{\varphi}_c = -46$ which one can see in T 2.1). In this direction this treatment could be *improved.* It is obviously better to take at least $|\tilde{\varphi}^\circ|$ instead $\tilde{\varphi}_c$ for $|\tilde{\varphi}_c| > |\varphi^\circ|$.

— The tables T 4.2 and T 2.1 show that the „direct" method (II) is better for case „C" where there is a high doping level. Even method (IV) is not inaccurate, which can be *explained* by the fact that the choice $\tilde{\varphi}_1 = -\tilde{\varphi}_N$ is now adequate (see $\tilde{\varphi}_{II}$ in Figure 4.1c).

The general conclusion could be, that one can calculate *far more accurately* the point of inversion by x_{iIV} and x_{iIV_e}, i.e. using *our relations* (the error is less than 15%), than using the „classical" relation for $^\circ x_i$ (last one in TB.e) based on the „space charge" approximation.

4.2. — The temperature dependence of the space charge region width

In Section 2.4 we obtained interesting results, with the theoretical treatment, that the dependence $x_{P,N}(T)$ and $\Lambda_\varrho(T)$ have *maximum* values for a temperature Θ_{NU}, graded and abrupt transition. These results are shown in Fig. 2.4.

Aiming at the verification of these results we have made the experimental studies on the following bases:

— the lowest temperature which one can obtain in our laboratory*) is that of liquid nitrogen ($\approx 77°K$);

— this means that the transitions should be abrupt and with the concentration of impurity as high as possible (see T 2.2);

— for this puprose, the best technique is alloying, which is the easiest to make in germanium.

*) Laboratory for Semiconductors, Electrotechnical Faculty — Beograd, and Laboratory of Institute of Physics, Beograd.

a. — The preparation of specimens. Measurements were carried out using the following P-N junctions:

(AD) — emitter's junction of a germanium transition which was made in IHTMI*) with *alloy-diffused* technique.

(Z) — Zener silicon diodes, produced in EI-Niš, with the breakdown voltage ≈ 20 V.

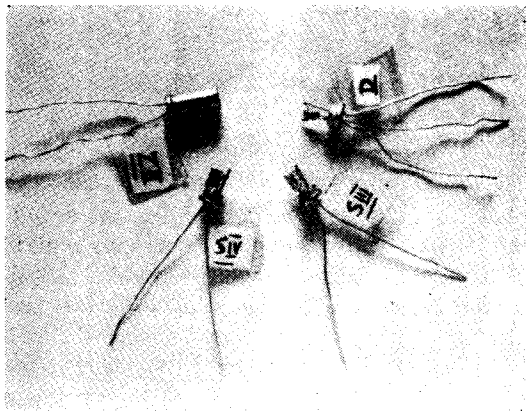

Fig. 4.2. — Photography of some specimen (S, Z, and AD) prepared for measurement.

(S) — Diodes specially made for this purpose*).

The last type (S) of transition was made, with alloying procedure, on N type Ge. The room temperature resistivity was about 0.01 Ωcm ($N \approx 2 \cdot 10^{17}$ cm^{-3}). The junctions were mounted on the standard headers and sealed with a proper resin to protect them from moisture (see Fig. 4.2).

We have chosen 5—6 specimens, from several dozen of all types, which have shown the smallest influence of the surface phenomenon, which was checked on a Tektronix 575 U-I tracer.

b. — The measuring method. A very well known and accurate resonance method has been chosen, as shown in Fig. 4.3. The elements were as follows:

— (S.G.) — High frequency generator (*Orion* 1163).

— (T) — High frequency transfomer with a resonant circuit in secondary.

— (Cs) — a standard variable capacitor (*Philips*, 60—360 pF and a *IEV* MA 2400 box × 100, × 1000 and × 10000 pF),

— (T.V.) — a tube voltmeter (*IEV* MA 3000),

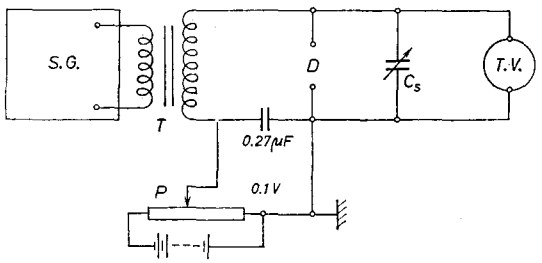

Fig. 4.3. — Circuit diagram for diode (D) capacity measurement.

— (P) — potentiometer for regulation of a DC bias for (D).

We have adjusted the oscillator circuit to be in resonance without diode, then connecting a diode the capacitance became bigger for C. Then one had to decrease the capacitance of C_s for the same amount without changing the frequency to obtain the resonance again which could be observed with an electronic voltmeter (T.V.).

To obtain a Q-factor big enough the resonant frequency had to be lower than 100—200 kHz, because the capacitance of transitions was several hund-

*) Institute for Chemical, Technology and Metallurgy investigations — Beograd, Laboratory for Semiconductors and special materials. The junstions were made by Ing. D. Đorđević.

reds pF for $U_{rev} = -0.1$ V. The capacitance of all diodes was measured for so small reverse voltage (0.1 V), because then the changes of temperature of x_{NP} were higher as follows from eqn. (2.31b). Really, for an abrupt transition $\delta_0(x_{N,P}) = 0$ and

$$x'_{NU} \equiv \frac{dx_{NU}}{dkT} = \frac{\tilde{\psi}_N - \tanh\psi_N \cdot \left(\frac{3}{2} + \frac{\tilde{W}_{gU}}{2}\right)}{d\psi_N/dx_N} = \ldots = F^2 \frac{\sqrt{1+\eta_{NU}^{-2}}\sinh^{-1}\eta_{NU} - \left(\frac{3}{2} + \frac{\tilde{W}_{gU}}{2}\right)}{x_{NU}},$$

(4.6a) $$F^2 \equiv \frac{2\varepsilon\varepsilon_0 P}{e^2 N(P+N)}; \quad \eta_{NU} = \frac{N}{2n_{iU}}, \quad n_{iU} = n_i\beta = n_i\, e^{-\frac{e\tilde{U}'}{2}}, \quad W_{gU} = W_g - eU',$$

since $\frac{d\psi_N}{dx_N} = \frac{x_{NU}\cdot N}{F^2(4n_{iU}^2 + N^2)^{1/2}}$, using (2.31) and (2.17b) with $\delta W_f^\circ = 0$. From (4.6a) one can see that x'_{NU} depends on U *mostly* thoughout the numerator x_{NU}, which could be seen even better if one assumes that $\eta_{NU} \gg 1$ (this is surely true for $U = U_{rev}$):

(4.6b) $$x'_{NU} \approx F^2 \frac{1\cdot\sinh^{-1}\eta_{NU} - (3+\tilde{W}_{gU})/2}{x_{NU}} =$$

$$= F^2 \frac{\ln 2\eta_{NU} - (3+\tilde{W}_{gU})/2}{x_{NU}} = F^2 \frac{\ln[N/A\,(e\,kT)^{3/2}]}{x_{NU}}.$$

When U increases — decreasing $|U_{rev}|$, X'_{NU} increases its absolute values similar to $\frac{1}{x_{NU}} \sim (U_D - U')^{-1/2}$ for $T = C^{te}$.

Diodes were not directly polarized, because a diffusion capacitance, which is more difficult to remove, would exist then.

The transition capacitance of the specimens was measured at four temperatures in a cryo-apparatus at about:

77°K — in liquid nitrogen,
195°K — dry ice CO_2 in acetone,
295°K — at room temperature and
370°K — about 95°C in a thermostate furnace (or 335 and 350°K, see T 4.3 too).

We need not verify *the accuracy* of the measurements because it was higher than necessary.

c. — Results. The measured values of the junction capacitances are given in T 4.3 as a function of temperature (for symbols S, Z and AD see the previous text). Special attention was paid to S — specimens, which were divided into three groups I, II and III with similar variations of $C(amb)/C(T)$ [$C(amb)$-capacitance at 295°K]. The dependence $C(T)$ for every specimen was measured for each temperature twice during decreasing and increasing T, so that the values in T 4.3 are *average* values (the dispersion was small-about $1^0/_{00}$ for all T, except for 370°K where it was smaller than 2%). Only diode Z II had abnormal behavior; its capacitance decreased with increasing T; it is obvious that temperature T was not low enough so that x_{NU} became

positive; with additional investigations it was shown that Z II had a pinch-through effect, so that its behaviour was not studied more.

C (T) dependence of junctions [pF] T4.3

T [°K] / The specimen	77	195	295	370 *)
S I	761	831	927	1140
S II	635	697	777	945
S III	801	881	977	1180
Z I	155	178	220	245
Z II	99	66	60	54
Z III	175	190	225	250
AD	11	15	30	69

*) For Z I and AD —335°K, but for Z III —350°K.

Because all transitions are abrupt and very asymmetrical ($P \gg N$, $\Lambda_\rho \approx x_N$) one can write, using a known axpression for $C(x_N)$ (see for instance [15]):

$$(4.6b)\quad C(T) = \varepsilon(T)\,\varepsilon_0 \frac{S}{\Lambda_\rho(T)} \approx \varepsilon\varepsilon_0 \frac{S}{x_{NU}} \therefore$$

$$\frac{\Lambda_\rho(T)}{\Lambda_\rho(amb)} \approx \frac{x_{NU}(T)}{x_{NU}(amb)} \approx \frac{C(amb)}{C(T)} \equiv y(T),$$

where room temperature values were taken as standard ones, („*amb*"— 295°K) and where we assumed that the changes of $\varepsilon(T)$ were neglectingly small (for $\alpha_\varepsilon = \delta\varepsilon/\varepsilon\delta T$ are given values of order $2 \cdot 10^{-4}$ 1/°K—[27], while from T 4.3 one can see that $|\alpha_c| = \frac{\delta C}{C\delta T}$ is at least of order 10^{-3} 1/°K).

Using the data from T 4.3 we calculated $y(T) \sim x_{NU}(T)$ and the results are given in Fig. 4.4.

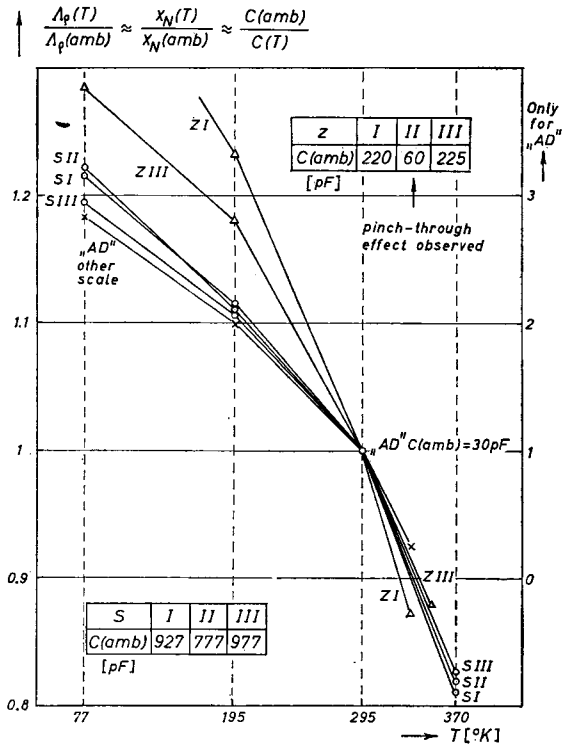

Fig. 4.4. — Dependences $x_{nU}(T) \sim y(T)$ according to (4.6c) and data from T 4.3. The scale for AD is on the right.

d. — The conclusion is as follows:

— the variations of $x_N(T)$ are *qualitatively* in *agreement* with theoretical consideration, x_N increases as T decreases, but more and more slowly. The maximum ($x_{NU\,max}$, see also Fig. 2.4 on the right), could not be obtained with our measurements, since judging by (2.31g) and T 2.2 the maximum is for $\Theta_{NU} \approx \Theta'_{NU}$ which should be $\leqslant$ 77/21 and 96/21 i.e. *below* 3.6 and 4.5°K ($21 = 100^{2/3}$), because in all specimens*) $N \leqslant 2 \cdot 10^{17}$ cm^{-3}. However a *tendency* towards maximum is obvious.

— using (4.6b) and F^2 from (4.6a) one can approximately write

$$x'_{NU} = \frac{2\varepsilon\varepsilon_0}{e^2 N x_{NU}} \ln \frac{N}{A(\mathrm{e}\,kT)^{3/2}} = \frac{2}{Se^2 N} C(T) \ln \frac{N}{A(\mathrm{e}\,kT)^{3/2}},$$

$$\therefore \frac{|x'_{NU}|}{x_{NU}\,(amb)} \equiv \left|\frac{dy}{dkT}\right| \equiv |y'| \sim \left|\ln \frac{A(\mathrm{e}\,kT)^{3/2}}{N}\right|^{**}, \quad \left(x_{NU} \sim \sqrt{\frac{\varepsilon\varepsilon_0}{N}}\right), \tag{4.6d}$$

where one can obtain x_{NU} and $C(T)$ using (2.17b) and (4.6c), respectively. Then one can see that with increasing N the slope $|y'|$ or $|x'_{NU}|$ will *decrease* for $T > \Theta_{NU}$ (for U and $T = C^{te}$). Our measurements show this and the agreement is not only qualitative but semi-qualitative too: the data from Figure 4.4 (say for 295°K) give

$$\frac{y'_{ZI}}{y'_{SI}} \approx \frac{2.1 \cdot 10^{-3}}{1.75 \cdot 10^{-3}} = 1.2 \tag{4.6e}$$

while with (4.6d) this ratio is 7.56/6.5 = 1.17, which is greater than one. This is in agreement with $N_Z < N_S$ too (A's are not very different for Ge and Si; using the equation 2.31g, they are $75 \cdot 10^{20}$ and $60 \cdot 10^{20}$ cm^{-3}/eV$^{3/2}$ respectively). One can obtain the same for (AD)—(S) but here it is not possible to make a quantitative comparison using (4.6d) since AD is a diffusion transition $\left(\text{with an „opposite" gradient: } \frac{dN}{dx} < 0\right)$ and (4.6d) is not valid even approximatively.

One can consider this effect experimentally almost proved: $\Lambda_p(T)$ in Figure 4.4 shows a clear tendency towards a maximum with decreasing T. It is possible to say that we have not used a very low temperature here, but later we have proved this effect *directly* (in [44]), as was said in Chapter 2, not only for Si and Ge transitions but for GaAs transitions too.

4.3. The verification of the relations for ionization coefficients

In Section 3.2b, using the theoretical observations of other authors, we obtained the relation (3.17b) for ionization coefficient α (preferably $y \equiv \alpha\Lambda$*) and we have applied it on the phenomena in P-N transitions. Now one should verify its quantitative side which will be done here for Ge and Si, although this treatment can be used for other materials too.

---

*) According to [28] for Si the breakdown voltage U_b is 20 V which corresponds to $N \approx 6.10^{16}$ sm^{-3}. Measuring U_b (on the tracer) of our „AD" and „S" transition one can obtain 6V and 4.5 V, respectively, which gives, by [28], $N = 10^{16}$ and 10^{17} cm^{-3}, respectively.

**) More exact calculation [21] yields $|y'| \sim \ln \frac{A(\mathrm{e}kT)^{3/2}}{\sqrt{PN}}$.

4.3.1. — The applied procedure

a. — Introduction

One should know the following *parameters* to be able to find out quantitative ionization coefficient as a function of the field K, for *every type* of carriers and for a *particular* semiconductor:

A — a dimensionless constant independent of the type of materials and carriers ($A = 3$),

B — is almost the same like A, but it *should depend* on the probability of ionization i.e. type of material and carriers. This could be found using (3.17a), (3.15a and c) and (3.14b).

W^* and $\hbar\omega_0$ — the ionization threshold of the atom and the energy of optical phonons, or parameter $\psi = k_\psi \frac{\hbar\omega_0}{C_0 W^*}$, $(k_\psi = C_0)$;

$\Lambda^* = \frac{\Lambda_0}{k_x}$ — the characteristic length.

W^* and Λ^* are parameters which *depend* on the type of carriers and materials; $\hbar\omega_0$ depends *only* on the material type.

We should note that one should know the parameters $\hbar\omega_0$ and W^* *separately* because they take part not only in $\psi = \hbar\omega_0/W^*$ but also in $x = \frac{W^*}{ek\Lambda^*}$.

b. Method of calculating the parameters

Taking into account the experimental function $\alpha_t(K)$ for a particular material and the type of carriers, the ideal treatment would be to change all *four* parameters, using a computer, and to try to find an optimal fitting with the function (3.17b). However the parameters $\hbar\omega_0$ and B have been experimentally obtained, although the various authors give different values for them (see T 4.4) so that the method mentioned could only be formally correct.

Because of that, we have proceeded as follows:

1° — One can assume the parameters B and $\hbar\omega_0$;

2° — One can calculate W^* and Λ^* from two sets of values (α_1, K_1) and (α_2, K_2) for α and K obtained from a known curve $\alpha_t(K)$ and our equation (3.17b). The calculation cannot be analytical, because the equations are transcendental and we obtain our solutions using a graphical method drawing W^* as a function of Λ^*, using (3.17b), with parameters α_1, K_1, B and $\hbar\omega_0$ and α_2, K_2, B and $\hbar\omega_0$ (the solution for the right values of W^* and Λ^* is obviously at the intersection of two curves*).

3° — With the above obtained values for the parameters, we have drawn a curve $\alpha_r(K)$ which is *compared* then with a known one (α_t); if the agreement was not satisfactory (although α_r and α_t must cut each other at two points), we repeated the whole process for another set of parameters.

c. — We could not verify *the exactness* of such a procedure so that results obtained are more semiquantitative than quantitative, since even the „exact" dependences $\alpha_t(K)$ are not very raliable.

*) Sometimes we have used an exact diagram from [39] instead of (3.17b) but then we had to use a graphical solution for obtaining the corresponding values W^*—λ^*. A. Mutavdžić helped us in performing this calculation (Electrotechnical Faculty — Beograd).

4.3.2. — The results and the comments

a. — The starting data are adopted from the literature whose values are given in T4.4. One should remark, once again, that those data *are not quite* exact, but they may be regarded as starting values. Some authors obtained those data finding the best fitting between the experimental results and the theory.

The source data for parameters and $\alpha_t(K)$ T 4.4

What[1]	For material and the carriers[2] Germanium n	Germanium p	Silicon n	Silicon p	Using: (remark)
Λ^* [Å]	70	100	70	100	[28] — 1957.
	130		—		[28] — 1955.
	—		160 — 200 — 260		[40]
W^* [eV]	1.5	3.5	1.5	3.5	[28] — 1957.
	1.5		—		[28] — 1955.
	—		1.5 — 2.3 — 3.5		[40]
$\hbar\omega_0$ [eV]	$25 \cdot 10^{-3}$		$51.5 \cdot 10^{-3}$		[31] — 1958.
	$36 \cdot 2 \cdot 10^{-3}$		—		($\Theta_0 = 420^0$ K)
	—		0.1		[33], ($\Theta_0 = 1200^0$ K)
$\alpha_t(K)$	B. Vul *et al.* [25].		S. Miller [28] — 1957 Figure 3.		see Fig. 4.5 too

[1] For the symbols and the names see the previous section (4.3.1) too.
[2] „n" for electrons; „p" for holes.

b. — The calculation and results

Using the mentioned data, and the treatment from section 4.3.1b, in [21] we have made a calculation for four cases: n-Ge, p-Ge, n-Si and p-Si.

We used such *a method* that at the *beginning*, procedure (I) was followed, using the data *underlined* in T4.4 and $B = 0.4$. In procedure (II) and (III), we assumed two parameters (B and $\hbar\omega_0$) and the other two were *calculated* (Λ^* and W^*, see paragraph 2° in Section 4.3.1b); the actual calculation for the last two parameters will not be given here, because it is a long procedure, but not very important.

c. — The comment and the conclusion

1° — Comparing results with the exact ones [$\alpha_t(K)$] in Figure 4.5 one can easily see that the *best agreement* can be obtained:

— for *electrons in germanium* using procedure (I), i.e. taking parameters from the literature (see T4.4 too) and taking $B = 0.4$, but not our calculation procedure (II) and (III). One can explain this that the movement of electrons is the best known and treated of all carriers, and from all materials germanium

is the best known too, so that the exactness of the data given in T4.4 is large just in this case.

— For *holes in germanium* — using method (III). This *confirms* once again our remark that B cannot be the same for electrons and holes.

— For *electron in Si* — using method (III). This shows that B must depend on the type of material, and also Λ^* which is not taken into consideration in the literature. Besides that, it seems that even the value $\hbar\omega_0 = 51.5\ 10^{-3}$ eV (Θ_{opt}—600° K) is not correct for Si (it is taken from [31]), and that a better one is about 0.1 eV.

— For the *holes in Si* the best agreement is obtained using method (III), and the comment may be the same as for n-Si.

In T4.5 we took parameter values which give the best agreement. The symbols are as before.

The Optimal Parameters T4.5

Parameter	Materials			
	Ge		Si	
	n	p	n	p
Λ^*[Å]	70	100	39	120
W^*[eV]	1.5	3.5	1.5	3.5
B	0.4	0.65	0.96	0.29
$\hbar\omega_0$ [eV]	$36.2 \cdot 10^{-3}$		0.1	
A	3			

We should emphasize that we involved *two more* parameters B and „Universal" A, besides obtaining more accurate values for parameters Λ^*, W^* and $\hbar\omega_0$. In connection with the above, relating B's and A with the basic quantities one could make a better analysis and discussion, which will be done elsewhere.

From the diagrams of figure 4.5 where function $\alpha_r(K)$ were drawn using (3.17b) for parameters given in T4.5, one can see the following:

— the agreement is better for larger K although it is quite good even for *decade lower values* of α:

— it is difficult to give an exact quantitative verification of our relation (3.17b) with parameters from T4.5, but one can easily see that it *does not give worse results* than the methods which have been used up to date (for instance [40], [28], [39] etc.), although it is much simpler and has the advantage that it has an analytical character;

— one can see, comparing α_r-the empirical relation (3.22a) (this is not given in figure 4.5) that the agreement between the real dependences α_t and ours is very bad especially for higher K where a mild „saturation" should exist. That is quite logical because (3.22a) is not justified conceptionally-physically. If one tried to find the relations in the shape of power function it is better to try to find them in the types of relation like (3.22c) which are the consequence of physically correct relations of type (3.17b).

3° — The *application* of our relation could be the following:

— obtaining the above mentioned simple but physically adequate relation $\alpha \sim K^r$,

— deriving the quantitative relation for *factors* of *multiplication* M_n and M_p (using eqn. 3.18g),

— the application of all the above mentioned on the *limiting regimes* (as in Section 3.2d): function ${}^0\mathcal{B}(U)$ using (3.22a), ${}^2\mathcal{B}(U)$ using (3.21b) and similar.

We are not going here to consider the applications, because that is outside the scope of this work. However, the previously mentioned verification is satisfactory, so *all the conclusions* and considerations from Section 3.2 *remain valid.*

However, one should remark that we have already considered the application of our relations (3.17b) and (3.22c) in Sec. 3.2d paragraph 1°. Deriving (from

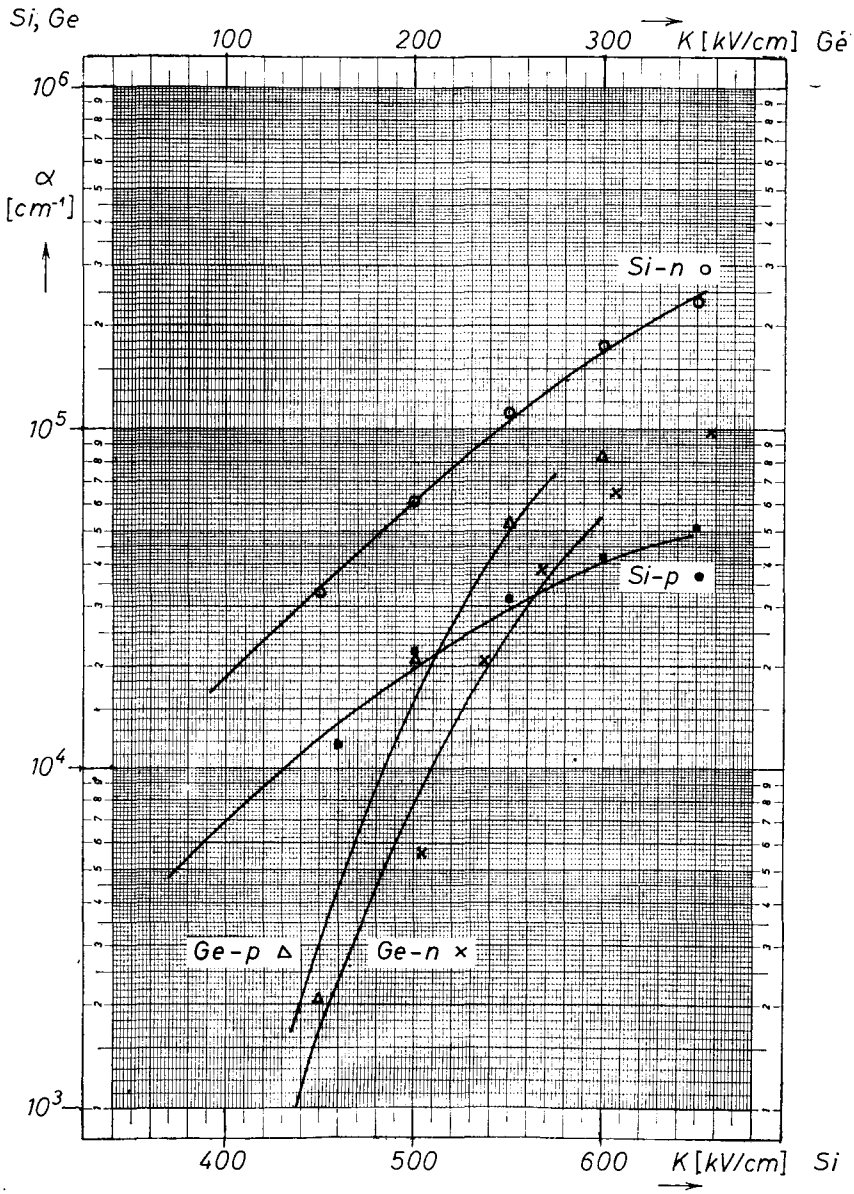

Fig. 4.5. — Ionization coeficients in dependence on the field:
— full lines — $\alpha_i(K)$ according to T 4.4,
— points (x, Δ, ○ and ●) without lines — according to T 4.5 and our relation (3.17b).

the mentioned relations) the relation (3.21d) *for the breakdown field* K_b and taking the parameters just from T4.5, for n-Ge we obtained an excellent agreement of calculated and experimental values ($U_b = 26.6$ V to 27 V): better than 1.5% for the breakdown voltage*).

4.4. — Comparison and checking of $I_3^*(\alpha^*)$ and $G(y)$ dependences

4.4.1. — Comparison of „recombination" integrals I_3^* and I_3

a. — **General.** In Section 3.1.a it was shown that *instead* of integral $I_3(\alpha)$ defined in [22] one *should put* our integral $I_3^*(\alpha^*)$ defined with (3.2c), into the relations for the recombination part of currents. For the case of an *abrupt* and very asymmetrical transition, which has not been treated in [22], we succeded in expressing I_3^* *using* tabulated function — the expression (3.5b). It is useful to know the derivative $I_3^{*\prime}$ of these integrals given with (3.5d).

b. — **The results** of the calculation are given in [21] while the diagrams in figure 4.6 are used for making the comparison.

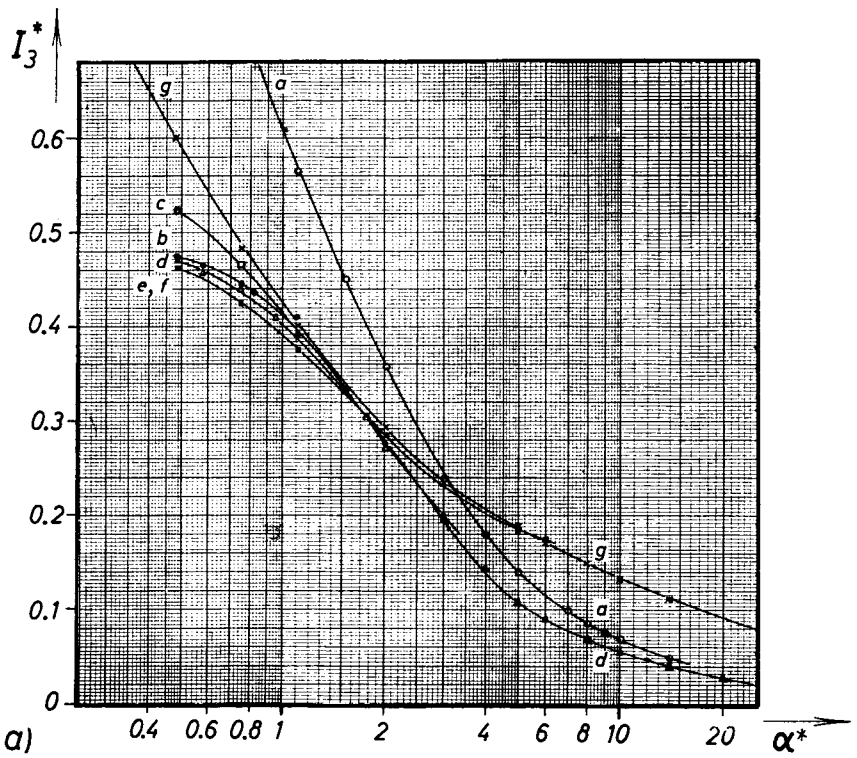

Fig. 4.6. — Graphic representation of „recombination" integrals I_3^* and their derivatives $I_3^{*\prime} = dI_3/d\alpha^*$ (*b*). For the signs (*a*, *b*, ...) see the text.

*) Although such a good agreement cannot be just a coincidence, one should not conclude that we have a pretension for the absolute quantitative accuracy of our relations. This question will be considered somewhere else (see for instance [46]).

The symbols are the following for both Figures: I_3^* and $I_3^{*\prime}$.

1° — *Symmetrical linear graded transition:*
(a) — according to [22] ($\alpha^* \to \alpha$ and $I_3^* \to I_3$);
(b) and (c) — using (3.4b and c): series expansion near $a=0$ and $a=2.095$;
(d) — graphical integration (this was done by M. Smiljanić and D. Slavić).

2° — *Extremely asymmetrical abrupt transition:*
(e) — according to (3.5b) and (3.5d);
(f) — graphical integration (M. S. and D. S. have done this).
(g) — like (e), but with $\Phi_{1,2,3,4}=1$, so that it is

$$\sum_1^4 \frac{(-1)^{n-1}}{(2n-1)^{1/2}} = 0.4919 \text{ and } I_3^*(\alpha^*) \approx \frac{0.418}{(\alpha^*)^{1/2}}, \quad \left(\alpha^* = \frac{e\tilde{U}_D - e\tilde{U}'}{2} + \frac{\delta \tilde{W}_f^0}{2}\right). \tag{4.7}$$

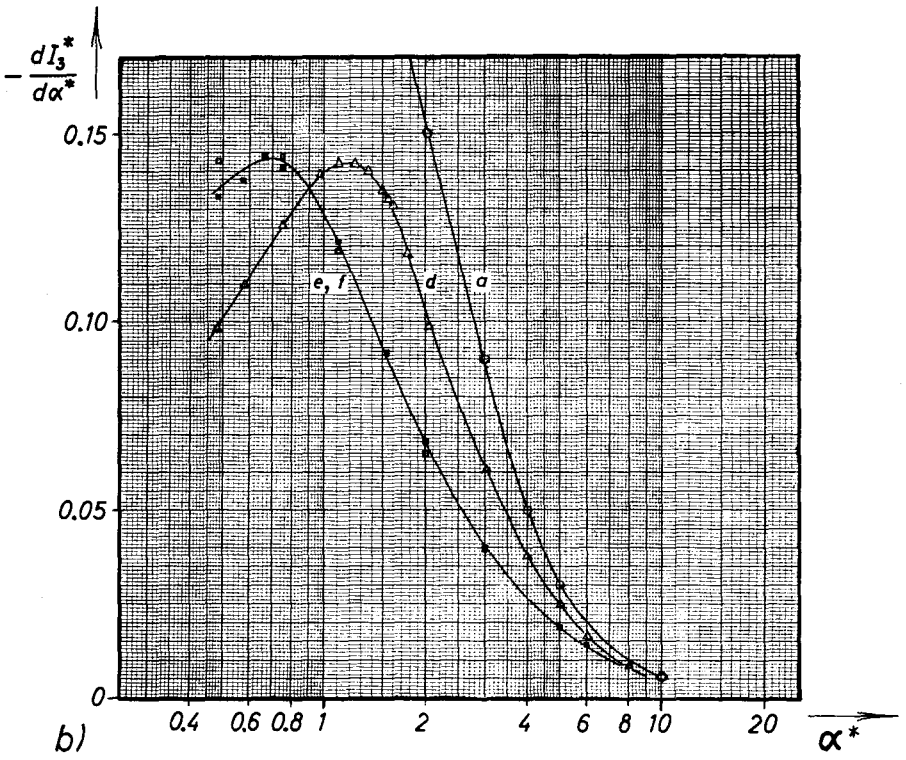

Fig. 4.6. b

3° — *Exceptionally* the symbol (a) in Figure 4.6b shows the results which we have obtained with the graphical method from the curve (a) in Figure 4.6a, since $\frac{dI_3}{d\alpha}$ has not been calculated in [22].

c. — The comment

1°. — For a *symmetrical graded* transition

— one can see an important departure of dependences (a) according to [22] from exact ones-(d), and not only for a *high* injection level ($\alpha^* \leqslant 1$) but also for .,medium" one where the disagreement is even 30%;

— the dependences (a) and (d) coincide only for $\alpha^* \geqslant 25$, i.e. for *very low* levels;

— our approximate relations (3.4b) — curve (b) and (3.4c) — curve (c) are in a good agreement*) with the exact one — (d) in their regions: $\alpha^* \leqslant 1$ and for α^* between 1 and 3, respectively.

2°. — We cannot compare the results, for an *asymmetrical abrupt* transition, with the results from the other authors because, as far as we know, there are none, but we can say

— that our expressions (3.5c) and (3.5d) are *quite axact***), because they are in agreement with the results of the graphical integration — with curve (f),

— that one can use even the *approximate relation* (4.7) for higher injection levels: for $\alpha^* \geqslant 0.8$ the error is smaller than 10%, while for $\alpha^* > 4$ it is *very small.*

3°. — *The General Conclusion* would be that *our results:* analytical, calculated and graphical, could be *reliably used* for quantitative analysis of the respective processes in the space charge region of P-N transitions.

4.4.2. — The verification of the validity of the relation (3.11b) for $G(y)$

a. — The method of verification of the relation (3.11b) derived for a limiting regime in transition, which gives an approximate solution of the differential equation (3.10b), is based here on the comparison of the numerical results obtained from relation $G(y)$ with more *exact* dependences obtained in some other way.

Although this method is not general, as for instance that used in Section 4.1.1, it could be accepted for a *rough estimation* of the quality of the expression (3.11b) connected to the departure of the regime from a currentless state***) (with a parameter A which is different from zero).

Among the exact dependences we know only one, that from [34]; that is relation $G(y)$ obtained with numerical integration for a case of *direct* recombination (and many values of parameter A between $+50$ and -50): the recombination parameters in equations (3.10c) or (3.9e) being $a=0$ and $c=1$.

The ratio of mobilities $b=\mu_n/\mu_p$ was taken to be 1.5****) and we shall assume the same value in our calculations.

---

*) In Figure 4.6a the curve (b) is moved, on purpose, from (d).

**) With a further verification it was shown that the divergence of shapes (e) and (f) does not exist (Figure 4.6b, $\alpha^* \approx 0.6$).

***) For $A,J=0$ we have obtained an *exact* solution (3.11a).

****) In [34] it is assumed that $b=1.5$ for Ge, but now it is well known that b is about 2.

b. — The results of the calculations (for the current parameters A as in T 4.6) are given together with those from [34] in [21].

Values of parameters A requested T4.6

A	+12	+4	+2	+1	0	−1	−12	−50
A_1	+ 9	+3	+1.5	+0.75	0	−0.75	− 9	−37.5

The symbols „+" and „—" designate *direct* and *reverse* direction of currents, respectively.

One should note that we could not find quite exactly the values from [34], because we have read them from diagrams in log—log scale, since in [34] the numerical values were not given. So here is a possible error of about 2—3%.

c. — Discussion of results and conclusion

1°. — For $A=O=J$ the agreement is *excellent*, which one need not verify, since the relation (3.11a) is quite exact, as it is already mentioned.

2°. — For the same absolute values of the current parameter A (for instance ± 12, ± 1) *the discrepancies are larger* for a direct polarization ($A>0$) and that is less convenient for us. For $A=-50$ the largest error is only about 12% (for $y=0.05$).

3°. — If we mark our values with G_c and those exact ones from [34] with G_r, the *ratios* will be as in T 4.7

The ratio G_c/G_r according to [21], A—parameter T4.7

	G_c/G_r for values of A:							
y	+12	+4	+2	+1	0	−1	−12	−50
0.05	0.81	0.64	0.98	0.956	1.008	**0.95**	**0.925**	**0.875**
0.1	0.66	**0.55**	**0.86**	**0.905**	0.99	1.01	1.037	0.965
0.5	**0.487**	0.58	0.872	0.918	0.995	1.00	0.92	0.924
1	0.524	0.67	0.931	0.900	0.987	0.98	0.984	0.943
5	0.744	0.846	0.92	0.945	0.964	0.98	1.03	0.987
10	0.875	0.88	0.93	0.966	0.975	0.983	—	1.07

Analysing these results we made the following *conclusions*:

— the maximal *departures* (figures in bold type) increase with increasing $|A|$ (they increase less for $A<0$) and appear for smaller values of y, when A *decreases* (for instance between +12 and −50 and further).

— for a *reverse polarization*, except for $y=0.05$, we can say that the agreement is satisfactory, practically for *all* values of A. This is even more favorable because the inverse currents are much smaller than the direct ones,

so that the case $A=-50$ practically *cannot* happen (see the equation 3.10a for A, too);

— for $y\geqslant 5$, G practically does not depend on A; this is valid for asymmetrical transitions and the limiting regimes.

4°. — The case of *direct currents*, when the *largest departures* occur (for $A>0$), is of a special interest. We shall consider this problem in connection with its application to the *limiting regimes* in P-N transition. We shall estimate the value for A, trying to find the currents $J^{\circ}_{p,n}$ for $x=x^{\circ}_{N,P}$, because the currents are then almost pure diffusion. As it was said in Section 3.1 before, but for *symmetrical transitions*, we shall have, taking roughly that $D_{pN}=D_{nP}=D_0$ and $L_{pN0}=L_{nP0}=L_0$, with all the symbols as before:

$$J^{\circ}_{pN}=\nu eD_0|z^0|\,N=\nu eD_0 N\frac{|G^0|}{L_0}\approx J^{\circ}_{nP}\,.$$

(4.8a)
$$J\approx J^{\circ}_{pN}+J^{\circ}_{nP}=2\nu D_0 N\frac{|G^0|}{L_0}e \quad\text{and}$$

$$A\equiv\frac{JL_{pN0}}{eD_{pN}\cdot Nb}\approx\frac{2\nu}{b}G^0\leqslant\frac{2\nu}{b}y^0;\quad \nu\approx 1+\frac{y^0}{1+y^0},$$

where we have written the expression for ν taking into account $D_p(p)$ from (3.8d). The sign „less" i.e. $|G^0|\leqslant y^0$ can be explained with the fact (from [21]), that for $A>0$, $|G|$'s are *always smaller* (or slightly bigger) than the value of y, if $y\leqslant 1.5$.

Taking into account (2.16c) for limiting regime we have $y^0\leqslant 0.618$, that means that $\nu\leqslant 1.38$ and finally, assuming even $\nu=1.5$, it is

(4.8b)
$$A<\frac{2\cdot 1.5}{1.5}y=2y \quad\text{for}\quad y<1.5,\ (b\geqslant 1.5),$$

that means that for $y^0<1$ parameter A is not larger than 2, for limiting regimes of direct polarized P-N transitions.

Using the above speculation and T 4.7 (the framed part) one can make a table of the largest errors (T 4.8):

T4.8

y	0.05	0.1	0.5	1	1.5
$A\leqslant$	0.1	0.2	1	2	3
the error	0	1%	8%	7%	≈10%

5°. — *The Conclusions* are as follows:

— our relation (3.11b) *practically* always quite well interprets the dependence of gradient of concentration versus concentration of carriers for *reverse* polarization (for $|A|<50$),

— while for *direct* polarization it can be used well for solving the problem of limiting regimes of the symmetrical transitions: for $y=y^0<1.5$, *the error is smaller* than 10%,

— qualitatively, our relation *shows correctly* the shape $G(y)$ for inverse polarization in a much wider region of y's and A's.

We should remark that the above consideration has not an absolute rigorous character, because it is based on a comparison with the more accurate results for a special case—direct recombination ($a=0$). However if an agreement was obtained for $a=0$ there is no reason to say that there will not be an agreement for $a\neq 0$. Actually, in a later work [47] we have verified our relation for a general case and *indirect* recombination (for $a=0$; 0.25; 1 and 2.5) calculating G_τ on a computer (Elliot 803). The results are in agreement with the above conclusions.

One of the possible *applications* of our expression for $G(y)$ is in studying the „breakdown" in a direct direction, i.e. — negative resistance of S-type (like in [35] and [36]), which could be the subject of another work.

Chapter 5. — OUTLINE OF THE PRINCIPAL RESULTS

On the basis of the analysis in the previous chapters the following principal conclusions, both with respect to the general theory and the more specific results, can be drawn.

a. — General theory

1° — The expression for the *quasi-neutrality condition* (Eqn. 2.20 c) has been obtained, which is more general than the results so far derived in the sense that it takes into account not only the impurity concentration ($\mathcal{N}$) but also the influence of the space charge (ϱ). In fact, in all previous works it was tacitly assumed that the condition $\varrho\approx 0$ must lead to $\vec{\nabla}\varrho\approx 0$ and $\Delta\varrho\approx 0$ which is not true in all cases (see, for example Fig. 2.2). Besides. the relation (2.20 c) indicates that the quasi-neutrality condition is also dependent on the *polarization* (U).

2° — The *transport equations* (2.2—2.4, 2.7) are given in the form which is especially suitable for consideration of the limiting working conditions i.e. the high reverse polarization or large forward current. It has been stated explicitly that the carrier concentrations (n, p) and the transport coefficients (D, μ) both depend on the position (r) and the field intensity (K) in the transition region, and that the lifetime (τ) is not always constant. In these equations the terms representing the „intrinsic" and „external" recombinations are separated.

3° — Special attention has been given to the *boundary conditions* and a new relation for the heterojunction has been derived (Eqn. 2.22). Different effective masses (m_A^0, m_A^∞, m_B^0, m_B^∞) have been introduced in this equation in order to comply with the fact the electron motions in various crystal lattices and in various states (in the solids A and B) are not the same. It is worth noting in passing that in the cited works the fact that these masses differ one from the others has not been respected and that the same true mass (not even the effective one) has been used throughout, which, at least in our opinion, might be considered as a shortcoming from the point of view of the general conception.

The *relations* (2.16 a) and (2.16 b) for the *P-N transition*, which have been widely used in Chapter 3, do not seem to have been derived yet. They take into account the band edge correction quite accurately (Fig. 2.1) and therefore, they can be widely used when studying the effects of high injection levels and reverse polalization*.

4° — In Section 3.2 c, a new formulation of the *breakdown condition* has been given which, from the point of view of physical interpretation, seems to be advantageous when compared with those in current use ($U=U_b$—breakdown voltage for $I\to\infty$). Instead of using the assumption that the breakdown occurs when the current increases to very large values,

* *Note added in the proof.* There is an attempt in the paper by Nussbaum [48]—1969 to take also non-zero gradients of the quasi-Fermi levels into consideration when boundary conditions are examined.

accerding to the new formulation the only criterion for the breakdown condition is that an incremental change in applied voltage causes a large increase of the current $\left(\frac{dI}{dU} \to \infty\right)$, while the current itself may have a finite value.

b. — More specific results

1° — It has been defined quantitatively *the least transition slope* ($\mathcal{S}_{0\,min}$ — Sec. 2.3) of the graded transition necessary to enable that the depletion region penetrates the region of the constant impurity concentration. The temperature dependence of the least transition slope has also been considered and shown that the shape of the function $\mathcal{S}_{0\,min} = f(T, U)$ exhibits a *minimum* (for $T = \Theta_{0U}$) which decreases with decreasing U (see Fig. 2.5). The smaller the values of U, the smaller the values of Θ_{0U} corresponding to the minimum of $\mathcal{S}_{0\,min}$ (these values of Θ_{0U} are extremely low — see Table 2.2).

2° — The *width of space charge region* has been studied in quantitative terms ($x_{N,P}$ or Λ_ρ, Section 2.4). It has been demonstrated that with *increasing temperature*, $x_{N,P}$ first increases, passes through a *maximum* at $T = \Theta_{N,PU}$, and then decreases in the region of very low temperatures (see Table T2,2 and Section 4.2).

It is interesting to note that *graded and abrupt junctions behave differently* when increasing reverse polarization, namely $\Theta_{N,PU}$ for graded junction decreases while for abrupt junctions increases with increasing reverse polarization (see Fig. 2.4).

This previously unknown phenomenon has been verified experimentally for abrupt junctions and it has been found a fairly good agreement between the measured results and the theory. (For complete proof see the author's paper [44]).

3° — Major attention has been given to the question of the *potential distribution* in the transition region and especially for the equilibrium states and abrupt transitions.

Our relations (2.22), which were derived by retaining the terms up the fourth power in the series expansion for $\cosh \tilde{\varphi}$, are more exact than all similar relations previously derived. As it has been shown in Section 4.1, they also offer the following advantages:

— They cover the *whole* P-*region* of the *asymmetrical* transitions ($P > N$) for any degree of asymmetry $\gamma = \frac{N}{P}$, with the error which is in all practical cases less than 6% as compared with the error of 22% for the „exponential" approximation. If, on the other hand, $N = 0$ (i. e. P-I transition) the error is even smaller than 0.5%. (This may be compared with 6.1% error obtained by the method outlined in reference [23]). Finally, for the case of a strong asymmetry*), the error was found to be less than 0.5%.

— In the *N-region* as well as in the whole transition region the error does not exceed the value given by the equation (4.1 d). This means that our approximation (2.22) provides satisfactory results up to the value $|\tilde{\alpha}| = |\tilde{\varphi}_{N,P} - \tilde{\varphi}_C| \approx 2 \div 3$ and, therefore, it is advantageous when compared with the „exponential" type of approximation at least for $|\delta\tilde{\alpha}| = |\delta\tilde{\varphi}_C| \approx 1$.

By comparing the values given by eqn. (2.22) with the exact values obtained in digital computer it has been found that the approximation is fairly satisfactory for lower doping levels (Fig. 4.1). In any case, it is superior to both the „exponential" type of approximation and the so-called Sh-approximation.

It must be noted, however, that the approximation (2.22) can hardly be used for the region II (Fig. 2.3b). For this region, some other approximations (φ_{II}, φ_{IIe} and φ_{el} — see Appendix B and Fig. 4.1) have been examined and found to be fairly good for smaller doping levels.

4° — Using the expressions for φ, the equations for obtaining the values of the *inversion points* have been derived and found, by numerical evaluations in Sec. 4.1.2, that for

*) It should also be emphasized that the expression (4.3f) yields the lowest degree of asymmetry γ_m corresponding to a maximum specified error (say 10%) in the P-region for a given doping level $\left(\frac{N}{n_i}\right)$ of the opposite (N) region.

smaller doping levels the best results are obtained with x_{iIVe}, while, for larger doping levels the equation for x_{iIV} is most satisfactory. (See Tables B.e and B.d.). The estimated error is less than 15% which compares advantageously with the results obtained with the Sh-approximation ($^{\circ}x_i$). In this way it has been proved that the real inversion planes are much closer to the planes of the physical transition ($x=0$) than what might be expected if the Sh-approximation were used.

5° — A suitable definition of the *transitional (intermediate) region* (Sec. 2.2d) has been introduced and the relation (2.22) employed to predict its width (Λ_{up}). It has been shown that Λ_{up} may be either smaller or larger than $\Lambda_p = x_N - x_P$ (space charge width) for higher temperatures and direct polarizations and for lower temperatures and reverse polarizations, respectively. On the other side, for medium temperature and $U \sim 0$, Λ_{up} is of the same order as Λ_p.

6° — A novel method is used in treating the influence of the *recombination processes* on the transition *susceptance* (capacitance). The derived expressions (2.30.c,d,e) are quite general and suitable for detailed analysis which, however, has not been accomplished in the proposed work. It follows from these equations that the transition behaves as a pure reactance only in the low frequency range and for a sufficiently low recombination, while the recombination processes must be taken into account in the general case. Both active and reactive components are present (G and $C\omega$) and they are dependent upon frequency, recombination mechanism and the transition profile.

7° — Special attention has been paid to the *limiting regimes* including *avalanche ionization* phenomenon. The following major results hav been reached:

— The previously known relation (in [31]) for the coefficients of impact ionization ($\alpha_{n,p}$) has been conveniently rearranged by introducing the *parameter A* (which does not depend on the type of semiconductor material) and the parameters B (which are dependent on the material used), apart from the known parameters (Λ^*, W^* and $\hbar\omega_0$). In this way a more suitable form of equation for comparison with the exact numerical results has been obtained (Eqn. 3.17b).

The *argreement* of the results for $\alpha(K)$ obtained by the expression (3.17b) with the measured values (for both Ge and Si) can be considered satisfactory to some degree which is, at least, much better than if the empirical relation $\alpha \sim K^r$ is used.

— On the basis of the foregoing analysis, the *functions* ${}^{2}\mathcal{B}(U)$ are introduced (Sec. 3.2b). They enable a better insight into the validity of different expressions for the approximation of $\alpha(K)$ and also for the *breakdown criterion*, i.e. the breakdown*) field (K_b). Each parameter in this equation can be given a clear physical interpretation which is not always the case with the previously derived expressions (see, for example [42]). Besides, the equation (3. 17b) also indicates that the *empirical* expressions for $M_{n,p}$, as for example (3.22e), in which the exponents $n_{p,n}$ are assumed to be constant, are not adequate.

8° — The problems relating to the *limiting regimes for direct polarization* have been considered in more detail and many interpretations of the phenomena involved have been improved, such as:

— A more precise definition of the „recombination" integral I_3^* (Sec. 3.1a) has been introduced leading to more exact results when compared to those in current use (Fig. 4.6). It enables to establish in more quantitative terms the physical fact that, in each transition, with increasing the direct polarization, the *ratio* of the recombination and diffusion currents *decreases steadily*.

— Using the elementary functions, the *analytical solution* for $G=f(y)$, G being the generalized *concentration gradient* and y — normalized *carrier concentration*, has been obtained. It is interesting to note that, though this function has been derived by using a mathematical treatment which can hardly be considered quite justified, it yields *quite correct* results for reverse polarizations, while for direct polarizations these solutions can be successfully applied to solving the problems of the limiting regimes (for $y = y^{\circ} < 1.5$, the error is less than 10%; see Sec. 4.4.2, and also our more recent Reference [47].

*) This problem has been thoroughly treated for $U_b(N)$ in the Reference [46].

APPENDIX A — List of Symbols*)

A — Constant
a — Lattice Constant; Length
B — Constant; number; $B_{c,v}$ Effective Density of States
b — Constant; Ratio of Mobilities μ_n/μ_p
C — Capacitance; Constant
c — Velocity of Light
D — Diffusion Coefficient; Electric Induction
d — Interatomic Distance; Diameter
E — Energy (known with accuracy of one constant); E_f — Electrochemical potential, Fermi Level
e — Magnitude of Electronic Charge — **Absolute** Value
F — Force; Function
f — Probability Distributon Function; Function; Frequency
G — Electric Conductance ($1/R$); Total Generation (current); generalized concentration gradient
g — Generation Rate
h — Plank's Constant ($h = 2\pi h$)
I — Integral; Intensity (of el. current)
i — Number; „i-th"
J — Electric Current Density**)
j — Imaginary Unit ($\sqrt{-1}$); Number; „j-th"
K — Electric Field Intensity*)
k — Bolzmann's Constant (as a difference from k — wave number)
l — Length; number; „l-th"
M — A certain function or an abbreviation for various expressions; Avalanche multiplication factor
m — Mass (of electron if it is without subscript) m^* — Effective mass of carriers (or $m_{n,p}$)
N — Total donor concentration; number; $\mathcal{N} = \dfrac{N-P}{n_i}$ dimensionless impurity concentration
n — Electron concentration in conduction band (n_o — equilibrium) n_i — Intrinsic concentration; number
P — Acceptor impurity concentration; probability; power
p — Hole concentration (p_o — equilibrium); number
Q — Coefficient; Heat quantity
q — Electric charge
R — Electric resistance ($1/G$); Radius
r — Radius, Position vector; number, coefficient; Recombination Rate
S — Area (Surface); Saturation (in subscript)
$\mathcal{S} = |\nabla N|$ — Slope of Impurity Transition Profile
s — „s-th", of „s's', kinds (in subscript)
T — Absolute Temperature
t — Time

*) Only the symbols that are found throughout the text are given here, while the „local" symbols are left out (but they are explained in their sections), although they may coincide with the general ones.

The *MKSA* system is used and the relations are noted in a rationalized form. **Electric field** *is denoted with* K, not with E or $\mathcal{E}$.

**) Only „current" is frequently written down instead of this full term, (we think that there will not be any ambiguity).

U — Voltage*) (pot. difference), U' — Voltage drop at transition itself
u — Number; Coefficient
V — Volume; Potential energy
v — Velocity (v_d — drift velocity, v_Φ — Velocity of sound)
W— Energy difference (kinetic energy); $W_f=(E_f-E_c)$ — chemical potential — *Relative* Fermi level (E_c is ref. point)
w — Transition rate probability
X — Reactance, x — coordinate
Y — Admittance; $y=\frac{P}{N}$ — Normalized hole concentration
Z — Impedance; $z=y'=p'/N$ — Normalized gradient concentration
α — coefficient (of temperature); Carrier ionization coefficient
α,β,γ — Current gain coefficients; angles
γ — Ratio (need not be dimensionless)
Δ — Laplacian
δ — Finite difference sign; Delta-function
ε — Relative dielectric constant (ε_0 — absolute — for the free space)
ζ — Coefficient
η — Coefficient; some function
Θ — Characteristic (absolute) temperature
θ — Angle (of scattering)
k — Reserve symbol
Λ — Characteristic length ($\Lambda_{n,p}$ — mean free path of carriers; Λ_ρ — barriers)
μ — Mobility: $\mu_f=C^{te}-E_f$
ν — Frequency; Coefficient
ξ — Coefficient; Function
Π — Product symbol
ρ — Space charge density; density
σ — Electric conductivity (abbreviated — „conductivity")
τ — Time interval (lifetime; relaxation time)
$d\tau$ — Element of volume
Φ — Some (potential) function; flux; Work function
φ — (Electric) potential: $\varphi(=)\ E$; $\varphi=-e\Upsilon+c^{te}$
χ — See eqns. (2.16)
Υ — Electrostatic potential
Ω — Solid angle; Elementary cell volume
ω — Angular frequency

Some symbols for subscripts

a — acceptor
b — breakdown
c — conduction band
D — Debye's; diffusion
d — donor
e — electronic
f — Fermi's
g — gap
i — intrinsic
rev — reverse polarization
dir — direct polar.
I — ionic, ionization
n — conduction electron
N— Region with donors dominating
o — equilibrium; initial; zero
p — of holes
P — Region with acceptors dominating
S — surface; saturation
v — valence band
ρ — space charge

Other letters (or these letters) have the same meaning as the quantity symbol.

*) In order to avoid ambiguity, we have chosen the symbols for electric field and voltage less frequently used: K and U (not E and V).

Abbreviations

$A_{n,p}$ or $A_{n:p}$ means that *two quantities* are concerned — „A_n" and „A_p".
T — in text — „Table", e.g. Tl.2: „Table 1 in Chapt. 2", or
TB.a: „Table ***a*** in Appendix ***B***"
IR — intermediate region
NR — neutral region
SC — Space charge (region)

Appendix B — SOME EXPRESSIONS FOR $\tilde{\varphi}$

Expressions for $\tilde{\varphi}_{IV}$ according to eqn.(2.24b) ***TB.a***

$$\tilde{\varphi}_u - \tilde{\varphi}_{IV} = \frac{12\, C_2'\, e^{x/\Lambda'_{Di}}}{(C_2' \cdot e^{x/\Lambda'_{Di}} + \tanh \tilde{\varphi}_u)^2 - 3}; \quad \Lambda'_{Di} = \Lambda_{Di} / \sqrt{2\cosh\tilde{\varphi}_u},$$

$$C'_2 = \left(6 - \tilde{\alpha}_N \tanh\tilde{\varphi}_u + \sqrt{3\tilde{\alpha}_N{}^2 + 36 - 12\tilde{\alpha}_N \tanh\tilde{\varphi}_N}\right) \cdot (\tilde{\alpha}_N)^{-1} \cdot e^{-x_N/\Lambda'_{Di}},$$

$$\tilde{\alpha} = \tilde{\alpha}_N = \tilde{\varphi}_u = \tilde{\varphi}_N \text{ for } x = x_N = x_i = 0.$$

Expressions for $\tilde{\varphi}_{IIe}$ according to eqn. (2.24c) and for $\tilde{\varphi}_{II}$ — (2.24d) ***TB.b***

$$\frac{x - x_e^0}{\Lambda'_{Di}} = \sinh^{-1} \frac{\sqrt{2 C_N \exp \tilde{\varphi}_{IIe}}}{M_N}, \text{ where:}$$

$$C_N = \tilde{\varphi}_N \sinh\tilde{\varphi}_N - \cosh\tilde{\varphi}_N + \nu \geqslant 0; \quad 0 \leqslant \nu \leqslant 1 \text{ and } M_N = \sqrt{\frac{1}{2}\left((\tilde{\varphi}_N \tanh\tilde{\varphi}_N - 1 + \frac{\nu}{\cosh\tilde{\varphi}_N} - \right)\cdot}$$

$$\frac{x - x^0}{\Lambda'_{Di}} = \sinh^{-1} \frac{\tilde{\varphi}_{II} + (\tilde{\varphi}_N - 2\tanh\tilde{\varphi}_N)}{2M}, \text{ where } M = \sqrt{\tilde{\varphi}_N \tanh\tilde{\varphi}_N - \tanh^2\tilde{\varphi}_N}.$$

Expressions for $\tilde{\varphi}_{el}$ according to (2.24e) ***TB.c***

$$\frac{x - x_i}{\Lambda_{Di}} = \begin{cases} = -\dfrac{F\left[\sin^{-1}\left(-\tanh\dfrac{\tilde{\varphi}_{el}}{2}\right),\ k_1\right]}{(1+C_u)^{1/2}}, & \left(C_u > 1,\ \text{i.e.}\ \dfrac{N}{n_i} = 5.78\right) \\ = -\dfrac{\sqrt{2}}{2} \cdot F\left(\sin^{-1}\sqrt{\dfrac{\cosh\tilde{\varphi}_{el} - 1}{\cosh\tilde{\varphi}_{el} + C_u'}},\ k_2\right), & \left(-1 < C_u' < 1,\ \text{i.e.}\ \dfrac{N}{n_i} = 0.26\right); \end{cases}$$

$$C_u = \tilde{\varphi}_u \sinh\tilde{\varphi}_u - \cosh\tilde{\varphi}_u, \quad C_u' = \tilde{\varphi}_u \sinh\tilde{\varphi}_u - \cosh\tilde{\varphi}_u + \nu;$$

$$k_1 = \sqrt{\frac{C_u - 1}{C_u + 1}}, \quad k_2 = \sqrt{\frac{1 - C_u'}{2}}.$$

Expressions for x°, φ°, $\eta_{P,N}$ and inversion point x_{iII} and x_{iIV} ***TB.d***

$$x^\circ = \Lambda'_{Di} \cdot \sinh^{-1} \frac{(2\tanh\tilde{\varphi}_N - \tilde{\varphi}^\circ) - \tilde{\varphi}_N}{2M}; \quad \tilde{\varphi}^\circ = \frac{\cosh\tilde{\varphi}_P - \cosh\tilde{\varphi}_N + \tilde{\varphi}_P \eta_P + \tilde{\varphi}_N \eta_N}{\eta_P + \eta_N}.$$

$$\eta_P = \frac{P}{2n_i} = \sinh(-\tilde{\varphi}_P), \quad \eta_N = \frac{N}{2n_i} = \sinh\tilde{\varphi}_N;$$

$$\frac{x_{iII}}{\Lambda'_{Di}} = \sinh^{-1}\frac{\tilde{\varphi}_N - 2\tanh\tilde{\varphi}_N}{2M} + \frac{x^\circ}{\Lambda'_{Di}},$$

$$\frac{x_{iIV}}{\Lambda'_{Di}} = \sinh^{-1}\frac{6-\tilde{\varphi}_N\tanh\tilde{\varphi}_N}{\tilde{\varphi}_N\sqrt{3-\tanh^2\tilde{\varphi}_N}} + \sinh^{-1}\frac{\tilde{\varphi}_N - \tanh\tilde{\varphi}_N - 3/\tanh\tilde{\varphi}_N}{M} + \frac{x^\circ}{\Lambda'_{Di}}.$$

Expressions for x_e° (see TB.b), $\tilde{\varphi}_c$ and x_{iIIe}, x_{iIVe} and $^\circ x_i$ *TB.e*

$$x_e^\circ = -\Lambda'_{Di}\frac{\sinh^{-1}\sqrt{2C_N\exp\tilde{\varphi}^\circ}}{M}; \quad \tilde{\varphi}_c = \tilde{\varphi}_N - \frac{6}{\tanh\tilde{\varphi}_N}.$$

$$x_{iIIe} = \frac{\Lambda'_{Di}}{M_N}\sinh^{-1}\sqrt{2C_N} + x_e^\circ,$$

$$\frac{x_{iIVe}}{\Lambda'_{Di}} = \sinh^{-1}\frac{6-\tilde{\varphi}_N\tanh\tilde{\varphi}_N}{\tilde{\varphi}_N\sqrt{3-\tanh^2\tilde{\varphi}_N}} + \left(\sinh^{-1}\sqrt{2C_N\exp\tilde{\varphi}_c}\right)/M_N + \frac{x_e^\circ}{\Lambda'_{Di}};$$

$$^\circ x_i = \frac{P}{N+P}\Lambda_{Di}\left(\frac{2n_i}{PN}(P+N)\ln\frac{NP}{n_i^2}\right)^{1/2}\left[1-\left(\frac{1+N/P}{1+|\tilde{\varphi}_P|/\tilde{\varphi}_N}\right)^{1/2}\right].$$

Appendix C — REFERENCES*

[1*] — F. Braun. *Pogg. Ann.*, **153,** 556 (1874); ...

[2*] — A. E. Flowers. *Phys. Rev.*, **29,** 445 (1909); ...

[3*] — P. Collet. *Ann. Phys. (Paris)*, **15,** 265 (1921).

[4*] — G. Siemens und W. Demberg. *Z. f. Phys.*, **67,** 375 (1931).

[5*] — W. Schottky *et al., Phys. Z.*, **30,** 839 (1929); **32,** 833 (1931).

[6*] — J. Frenkel, A. Joffé. *Phys. Z. Sowietunion*, **1,** 60 (1932); ...

[7*] — Б. И. Давыдов. *ЖТФ*, **8,** 3 (1938); *ЖЭТФ*, **7,** 1069 (1937); ...

[8] — W. Shockley: *Electrons and Holes in Semiconductors.* New York, 1953.

[9] — А. И. Губанов: *Теория выпрямляющего действия полупроводников.* Москва, 1956.

[10] — H. K. Henisch: *Rectifying Semi-Conductor Contacts.* Oxford, 1957.

[11] — E. Spenke: *Electronische Halbleiter.* Berlin, 1955.

[12] — P. Aigrain, F. Englert: *Les Semiconducteurs.* Paris, 1958.

[13] — A. K. Jonscher: *Principles of Semiconductor Device Operation.* New York 1960.

[14] — Я. А. Федотов: *Основы физики полупроводниковых приборов.* Москва, 1964.

[15] — D. Tjapkin: *Electron Physics of Solids* (in Serbian), Book II. Belgrade, 1964.

[16] — J. Grosvalet. *Sol. State Electronics*, **6,** 63 (1963).

[17] — Procedings of „*Colloque international sur les dispositifs à semiconducteurs*" Paris, fév. 1961, *Vol. 1* et *2*. Ed. Chiron, Paris, 1961.

[18] — R. Hall. *J. Appl. Phys.*, **29,** 914 (1958).

[19] — Т. Харламова *et al. ФТТ*, **2,** 434; 1909 (1960).

*) Bibliography marked with [*] is not direct consulted, and that marked with [+] is added to this version in relation to [21].

If only the figure is cited in the text without the year, (e.g. [24]) it means that it is the first cited paper.

[20] — G. Greebe, *Philips Res. Rep. Suppl.*, № 1 (1963).

[21+] — D. Tjapkin: *Study of Current-Voltage Regimes in Semiconductor P-N Structures.* (Doctor's thesis, in Serbian). Electrical Engineering Faculty, Belgrade, 1966.

[22] — А. Д. Чевычелов, *ФТТ*, **1**, 1205 (1959).

[23] — Э. И. Адирович *et al.*, *ЖТФ*, **28**, 55 (1958); *Радиотехника и электроника*, **4**, 1708 (1959).

[24] — W. Shockley. *BSTJ*, **28**, 435 (1949); 335 (1949); **30**, 990 (1951); *PIRE*, **40**, 1374 (1952); *Sol. State Electronics*, **2**, 35 (1961); ...

[25] — Б. М. Вул (*et al.*). *ЖТФ*, **28**, 681 (1958); *ФТТ*, **2**, 2961 (1960); *ЖТФ*, **26**, 2403 (1956).

[26] — S. P. Morgan and F. M. Smits. *BSTJ*, **39**, 1573 (1960).

[27] — Л. С. Берман: *Нелинейная полупроводниковая емкость.* Москва, 1963.

[28] — S. L. Miller. *Phys. Rev.*, **99**, 1234 (1955) (for Ge); **105**, 1246 (1957) (for Si).

[29] — H. Brooks. *Adv. in Electronics and Electron Phys.*, **7**, 87 (1955).

[30] — E. M. Conwell, (V. Weisskopf). *Phys. Rev.*, **77**, 388 (1950); **88**, 1379 (1952); ...

[31] — Чуенков В. А. *Физика твердого тела*, Сборник статей, I и II, Изд. АН СССР, М., 1959; *ЖТФ*. **28**, 470 (1958); *ФТТ*, **2**, 799 (1960).

[32] — Chin-Tang Sah, R. Noyce, W. Shockley. *PIRE*, **45**, 1228 (1957).

[33] — J. Bok. *Ann. Radioélectr.* **15**, № 60 (1960).

[34] — W. van Roosbroeck. *BSTJ*, **29**, 560 (1950).

[35] — В. И. Стафеев. *ЖТФ*, **28**, 1631 (1958); ...

[36] — D. Tjapkin et Dj. Bugarinović. *Publ. de la Fac. El. de l'Université à Belgrade*, Sér. Math, et Phys., № 109 (1963).

[37] — F. van der Maesen *et al.*, *Philips Res. Repts.*, **17**, 479 (1962).

[38+] — Л. В. Кельдыш, *ЖЭТФ*, **37**, 713 (1959); **48**, 1692 (1965).

[39] — G. A. Baraff. *Phys. Rev.*, **128**, 2507 (1962).

[40] — P. A. Wolff. *Phys. Rev.* **95**, 1415 (1954); *Phys. Chem. Sol.*, **16**, 184 (1960).

[41] — A. G. Chynoweth. *Phys. Rev*, **109**; 1537 (1958); **118**, 2 (1960).

[42] — H. L. Armstrong. *J. Electron and Control*, **5**, 97 (1958).

[43] — K. G. Mc Kay (*et al.*). *Phys. Rev.*, **91**, 1079 (1953); **94**, 877 (1954).

[44+] — D. Tjapkin. *Proc. of the XIIth Conf. of Electronics, Telecom., Autom. and Nucl.* End., June 3—5, Belgrade 1968., p.p. 541 to 546 (in Serbian).

[45+] — M. Smiljanić. Reported at the *First Yug. Symposium of Solid State Physics*, Belgrade, June 27—29, 1968.

[46+] — As in [44] but ... *of XIth Conf.*, 1967; p.p. 397—405.

[47+] — M. Smiljanić and Tjapkin, *ibid.*, p.p. 451—457.

[48+] — A. Nussbaum. *Solid-St. Electron.* **12**, 177 (1969).