We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter 2

Physical Design Fundamentals of High-Performance Avalanche Heterophotodiodes with Separate Absorption and Multiplication Regions

Viacheslav Kholodnov and Mikhail Nikitin

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50778

1. Introduction

Minimal value of dark current in reverse biased $p-n$ junctions at avalanche breakdown is determined by interband tunneling. For example, tunnel component of dark current becomes dominant in reverse biased $p-n$ junctions formed in a number of semiconductor materials with relatively wide gap $E_g$ already at room temperature when bias $V_b$ is close to avalanche breakdown voltage $V_{BD}$ (Sze, 1981), (Tsang, 1981). The above statement is applicable, for example, to $p-n$ junctions formed in semiconductor structures based on ternary alloy $In_{0.53}Ga_{0.47}As$ which is one of the most important material for optical communication technology in wavelength range $\lambda$ up to 1.7 μm (Tsang, 1981), (Stillman, 1981), (Filachev et al, 2010), (Kim et al, 1981), (Forrest et al, 1983), (Tarof et al, 1990), (Ito et al, 1981). Significant decreasing of tunnel current can be achieved in avalanche photodiode (APD) formed on multilayer heterostructure (Fig. 1) with built-in $p-n$ junction when metallurgical boundary of $p-n$ junction ($x=0$) lies in wide-gap layer of heterostructure (Tsang, 1981), (Stillman, 1981), (Filachev et al, 2010), (Kim et al, 1981), (Forrest et al, 1983), (Tarof et al, 1990), (Clark et al, 2007), (Hayat & Ramirez, 2012), (Filachev et al, 2011). Design and specification of heterostructure for creation high performance APD must be such that in operation mode the following two conditions are satisfied. First, space charge region (SCR) penetrates into narrow-gap light absorbing layer (absorber) and second, due to decrease of electric field $E(x)$ into depth from $x=0$ (Fig. 1), process of avalanche multiplication of charge carriers could only develop in wide-gap layer. This concept is known as APD with separate absorption and multiplication regions (SAM-
APD). Suppression of tunnel current is caused by the fact that higher value of $E$ corresponds to wider gap $E_g$. Electric field in narrow-gap layer is not high enough to produce high tunnel current in this layer. Dark current component due to thermal generation of charge carriers in SCR (thermal generation current with density $J_G$) is proportional to intrinsic concentration of charge carriers $n_i \propto \exp(-E_g/2k_B T)$, here $k_B$ – Boltzmann constant, $T$ – temperature (Sze, 1981), (Stillman, 1981). Tunnel current density $J_T$ grows considerably stronger with narrowing $E_g$ than $n_i$ and depends weakly on $T$ (Stillman, 1981), (Burstein & Lundqvist, 1969). Therefore, component $J_T$ will prevail over $J_G$ in semiconductor structures with reasonably narrow gap $E_g$ even at room temperature. Another dark current component – diffusion-drift current caused by inflow of minority charge carriers into SCR from quasi-neutral regions of heterostructure is proportional to $n_i^2 \times N^{-1}$ (Sze, 1981), (Stillman, 1981) (where $N$ is dopant concentration). To eliminate it one side of $p-n$ junction is doped heavily and narrow-gap layer is grown on wide-gap isotype heavily doped substrate (Tsang, 1981). Thus heterostructure like as $p_{wg}^- n_{wg}^-$ is the most optimal, where subscript $wg$ means wide-gap and $ng$ – narrow-gap, properly. To ensure tunnel current’s density not exceeding preset value is important to know exactly allowable variation intervals of dopants concentrations and thicknesses of heterostructure’s layers. Thickness of narrow-gap layer $W_2$ is defined mainly by light absorption coefficient $\gamma$ and speed-of-response. But as it will be shown further tunnel current’s density depends strongly on thickness of wide-gap layer $W_1$ and dopant concentrations in wide-gap $N_1$ and narrow-gap $N_2$ layers. Approach to optimize SAM-APD structure was proposed in articles (Kim et al, 1981), (Forrest et al, 1983) (see also (Tsang, 1981)). Authors have developed diagram for physical design of SAM-APD based on heterostructure including $In_{0.53}Ga_{0.47}As$ layer. However, diagram is not enough informative, even incorrect significantly, and cannot be reliably used for determining allowable variation intervals of heterostructure’s parameters. The matter is that diagram was developed under assumption that when electric field $E(x)$ (see Fig. 1b) at metallurgical boundary of $p_{wg}^- n_{wg}^-$ junction $E(0) = E_1$ is higher than $4.5 \times 10^5$ V/cm then avalanche multiplication of charge carriers occurs in $InP$ layer where $p_{wg}^- n_{wg}^-$ junction lies at any dopants concentrations and thicknesses of heterostructure’s layers. However, electric field $E_1 = E_{1BD}$ at which avalanche breakdown of $p-n$ junction occurs depends on both doping and thicknesses of layers (Sze, 1981), (Tsang, 1981), (Osipov & Kholodnov, 1987), (Kholodnov, 1988), (Kholodnov, 1996-2), (Kholodnov, 1996-3), (Kholodnov, 1998), (Kholodnov & Kurochkin, 1998). As a consequence, avalanche multiplication of charge carriers in considered heterostructure can either does not occur at electric field value $E_1 = 4.5 \times 10^5$ V/cm or occurs in narrow-gap layer (Osipov & Kholodnov, 1987), (Osipov & Kholodnov, 1989). Value of electric field required to initialize avalanche multiplication of charge carriers can even exceed $E_{1BD}$ (Sze, 1981), (Osipov & Kholodnov, 1987), (Kholodnov, 1996-2), (Kholodnov, 1996-3), (Kholod-
nov, 1998), (Kholodnov & Kurochkin, 1998) that has physical meaning in the case of transient process only (Groves et al, 2005), (Kholodnov, 2009). Further, in development of diagram was assumed that maximal allowable value of electric field in absorber at hetero-interface with multiplication layer $E_2$ (see Fig. 1b) is equal to $1.5 \times 10^5$ V/cm. But tunnel current density $J_T$ in narrow-gap absorber $In_{0.53}Ga_{0.47}As$ (Osipov & Kholodnov, 1989) is much smaller at that value of electric field than density of thermal generation current $J_G$ which in the best samples of $InP-In_{0.53}Ga_{0.47}As-InP$ heterostructures (Tsang, 1981), (Tarof et al, 1990), (Braer et al, 1990) can be up to $10^{-6}$ A/cm$^2$. However, diagram does not take into account the fact that tunnel current in wide-gap multiplication layer can be much greater than in narrow-gap absorber (Osipov & Kholodnov, 1989). Therefore, total tunnel current can exceed thermal generation current.

In present chapter is done systematic analysis of interband tunnel current in avalanche heterophotodiode (AHPD) and its dependence on dopants concentrations $N_1$ in $n_{wg}$ wide-gap and $N_2$ in $n_{ng}$ narrow-gap layers of heterostructure and thicknesses $W_1$ and $W_2$, respectively (Fig. 1) and fundamental parameters of semiconductor materials also. Performance limits of AHPDs are analyzed (Kholodnov, 1996). Formula for quantum efficiency $\eta$ of heterostructure is derived taking into account multiple internal reflections from hetero-interfaces. Concentration-thickness nomograms were developed to determine allowable variation intervals of dopants concentrations and thicknesses of heterostructure layers in order to match preset noise density and avalanche multiplication gain of photocurrent. It was found that maximal possible AHPD’s speed-of-response depends on photocurrent’s gain due to avalanche multiplication, as it is well known and permissible noise density for preset value of photocurrent’s gain also. Detailed calculations for heterostructure $InP-In_{0.53}Ga_{0.47}As-InP$ are performed. The following values of fundamental parameters of $InP$ (I, Fig. 1) and $In_{0.53}Ga_{0.47}As$ (II, Fig. 1) materials (Tsang, 1981), (Stillman, 1981), (Kim et al, 1981), (Forrest et al, 1983), (Tarof et al, 1990), (Ito et al, 1981), (Braer et al, 1990), (Stillman et al, 1983), (Burkhard et al, 1982), (Casey & Panish, 1978) are used in calculations: band-gaps $E_{g1}=1.35$ eV and $E_{g2}=0.73$ eV; intrinsic charge carriers concentrations $n_i^{(1)}=10^8$ cm$^{-3}$ and $n_i^{(2)}=5.4 \times 10^{11}$ cm$^{-3}$; relative dielectric constants $\epsilon_1=12.4$ and $\epsilon_2=13.9$; light absorption coefficient in $In_{0.53}Ga_{0.47}As$ $\gamma=10^4$ cm$^{-1}$; specific effective masses $m^*_1=2m_e m_h/(m_e + m_h)$ of light carriers $m_i=0.06m_0$ and $m_2=0.045m_0$ where $m_0$ – free electron mass. The chapter material is presented in analytical form. For this purpose simple formulas for avalanche breakdown electric field $E_{BD}$ and voltage $V_{BD}$ of $p-n$ junction are derived taking into account finite thickness of layer. Analytical expression for exponent in well-known Miller’s relation was obtained (Sze, 1981), (Tsang, 1981), (Miller, 1955) which describes dependence of charge carriers’ avalanche multiplication factors on applied bias voltage $V_b$. It is shown in final section that Geiger mode (Groves et al, 2005) of APD operation can be described by elementary functions (Kholodnov, 2009).
2. Formulation of the problem: Basic relations

Let’s consider $p_{wg} - n_{wg} - n_{wg} - n_{wg}^*$ heterostructure at reverse bias $V_b$, sufficient to initialize avalanche multiplication of charge carries. This structure is basic for fabrication of AHPDs.

\[ M_n = M(-L_p), M_p = M(L_n), \bar{M}(-L_p, L_n) = \int_{-L_p}^{L_n} g(x)M(x)dx, \]

\[ M(x) = Y(x, -L_p)/(1-m), m(-L_p, L_n) = \int_{-L_p}^{L_n} \alpha(x)Y(x, -W_p)dx, Y(x, x_0) = \exp \left[ \int_{x_0}^{x} (\beta - \alpha)dx' \right] \]

(1)

(2)

can be determined, in principal, dependences of multiplication factors \( M \) in \( p-n \) structures on \( V_b \) where \( M_n \) and \( M_p \) – multiplication factors of electrons and holes inflow into space charge region (SCR); value of multiplication factor of charge carriers generated in SCR \( \bar{M} \) lies between \( M_n \) and \( M_p \); specific rate of charge carriers’ generation in SCR \( \bar{g} \) consists of dark \( g_d \) and photogenerated \( g_{ph} \) components; \( L_p \) and \( L_n \) – thicknesses of SCR in \( p \) and \( n \) sides of structure; \( \alpha(E) \) and \( \beta(E) = K(E) \times \alpha(E) \) – impact ionization coefficients of electrons \( \alpha(E) \) and holes \( \beta(E) \); \( E(x) \) – electric field. Let’s denote by \( N_{I1pt} \) dopant concentration \( N_i \) so that for \( N_i < N_{I1pt} \) “punch-through” (depletion) of \( n_{wg} \) layer occurs that means penetration of non-equilibrium SCR into \( n_{ng} \) layer (Fig. 1). Optical radiation passing through wide-gap window is absorbed in \( n_{ng} \) layer and generates electron-holes pairs in it. When \( N_i < N_{I1pt} \) then photo-holes appearing near \( n_{wg}/n_{ng} \) heterojunction \((x=W_1)\) are heated in electric field of non-equilibrium SCR and, at moderate discontinuities in valence band top \( E_v \) at \( x=W_1 \), photo-holes penetrate into \( n_{wg} \) layer (layer I) due to emission and tunneling. If \( W_1 \) is larger than some value \( W_{min} \)(Osipov & Kholodnov, 1989), which is calculated below, then avalanche multiplication of charge carriers occurs only in \( n_{wg} \) layer, i.e. photo-holes fly through whole region of multiplication. In this case photocurrent’s gain (Tsang, 1981), (Artsis & Kholodnov, 1984) \( M_{ph} = M_p \). Let \( p_{wg} \) layer is doped so heavy that avalanche multiplication of charge carriers in it can be neglected (Kholodnov, 1996-2), (Kholodnov & Kurochkin, 1998). Under these conditions thicknesses in relations (1) and (2) can be put \( L_p = 0 \) and \( L_n = W_1 \), i.e.

\[ M_{ph} = Y(W_1, 0)/(1-m(0,W_1)) \]

(3)

It is remarkable that responsivity \( S_I(\lambda) \) (where \( \lambda \) – is wavelength) of heterostructure increases dramatically once SCR reaches absorber \( n_{wg} \) (layer II on Fig. 1) and then depends weakly on bias \( V_b \) till avalanche breakdown voltage value \( V_{BD} \) (Stillman, 1981). This effect is caused by potential barrier for photo-holes on \( n_{wg}/n_{ng} \) heterojunction and heating of photo-holes in
electric field of non-equilibrium SCR. If losses due to recombination are negligible (Sze, 1981), (Tsang, 1981), (Stillman, 1981), (Forrest et al, 1983), (Stillman et al, 1983), (Ando et al, 1980), (Trommer, 1984), for example, at punch-through of absorber, then $S_I(\lambda)$ in operation mode is determined by well-known expression (Sze, 1981), (Tsang, 1981), (Stillman, 1981), (Filachev et al, 2011):

$$S_I(\lambda) = \eta(\lambda) \times \frac{\lambda}{12.4} \times M_{ph}$$

(4)

where $\lambda$ in $\mu$m and value of quantum efficiency $\eta$ is considered below. Photocurrent gaining and large drift velocity of charge carriers in SCR allow creating high-speed high-performance photo-receivers with APDs as sensitive elements (Sze, 1981), (Tsang, 1981), (Filachev et al, 2010), (Filachev et al, 2011), (Woul, 1980). Reason is high noise density of external electronics circuit at high frequencies or large leakage currents that results in decrease in Noise Equivalent Power (NEP) of photo-receiver with increase of $M_{ph}$ despite of growth APD’s noise-to-signal ratio (Tsang, 1981), (Filachev et al, 2011), (Woul, 1980), (McIntyre, 1966). Decrease in NEP takes place until $M_{ph}$ becomes higher then certain value $M_{ph}^{opt}$ above which noise of APD becomes dominant in photo-receiver (Sze, 1981), (Tsang, 1981), (Filachev et al, 2011), (Woul, 1980). Even at low leakage current and low noise density of external electronics circuit, avalanche multiplication of charge carriers may lead to degradation in NEP of photo-receiver due to decreasing tendency of signal-to-noise ratio dependence on APD’s $M_{ph}$ under certain conditions (Artsis & Kholodnov, 1984). Moreover, excess factor of avalanche noise (Tsang, 1981), (Filachev et al, 2011), (Woul, 1980), (McIntyre, 1966) may decrease with powering of avalanche process as, for example, in metal-dielectric-semiconductor avalanche structures, due to screening of electric field by free charge carriers (Kurochkin & Kholodnov 1999), (Kurochkin & Kholodnov 1999-2). Using results obtained in (Artsis & Kholodnov, 1984), (McIntyre, 1966), noise spectral density $S_N$ of $p_{n2}^{-} - n_{n2}^{-} - n_{wg}^{-} - n_{wg}^{+}$ heterostructure which performance is limited by tunnel current can be written as:

$$S_N = 2 \times q \times A_S \times M_{ph}^2 \times \sum_{i=1}^{2} T_{j,i}(V) F_{j,i}(M_{ph})$$

(5)

where $q$ – electron charge; $A_S$ – cross-section area of APD’s structure; $F_{j,i}(M_{ph})$ – effective noise factors (Artsis & Kholodnov, 1984) in wide-gap multiplication layer ($i=1$) and in absorber ($i=2$); $T_{j,i}(V)$ – densities of primary tunnel currents in those layers, i.e. tunnel currents which would exist in layers I and II in absence of multiplication of charge carriers due to avalanche impact generation. Comparison of two different APDs in order to determine which one is of better performance is reasonable only at same value of $M_{ph}$. Expression (5) shows, that for preset gain of photocurrent, noise density is determined by values of pri-
mary tunnel currents \( I_{T1} = J_{T1} \times A_S \) and \( I_{T2} = J_{T2} \times A_S \) (total primary tunnel current \( I_T = I_{T1} + I_{T2} \)). Distribution of electric field \( E(x) \) that should be known to calculate parameters (4) and (5) of AHPD is obtained from Poisson equation and in layers I and II is determined by expressions:

\[
E(x) = \frac{qN_1 x}{\varepsilon_0 \varepsilon_1} \times U_-(l_1 - x),
\]

\[
E(x) = \left[ E_2 - \frac{qN_2}{\varepsilon_0 \varepsilon_2} (x - W_1) \right] \times U_-(W_1 + l_2 - x),
\]

Where

\[
E_2 = \frac{\varepsilon_1}{\varepsilon_2} E_1 - \frac{qN_2 W_1}{\varepsilon_0 \varepsilon_2} \times U_-(l_1 - W_1),
\]

\[
I_i = \frac{qN_1}{qN_1} E_i \times U_-(W_i - l_i) + W_i \times U_-(l_i - W_i),
\]

\( U_-(x) \) and \( U_+(x) \) – asymmetric unit stepwise functions (Korn G. & Korn T., 2000), \( \varepsilon_0 \) – dielectric constant of vacuum, \( \varepsilon_1 \) and \( \varepsilon_2 \) – relative dielectric permittivity of \( n_{wg} \) and \( n_{ng} \) layers (Fig. 1).

3. Avalanche multiplication factors of charge carriers in p-n structures

3.1. Preliminary remarks: Avalanche breakdown field

For successful development of semiconductor devices using effects of impact ionization and avalanche multiplication of charge carriers is necessary to know dependences of avalanche multiplication factors \( M(V) \) of charge carriers in \( p-n \) structures on applied bias \( V_b \). We need to know among them dependence of avalanche breakdown voltage \( V_{BD} \) on parameters of \( p-n \) structure and distribution of electric field \( E(x) \) related to \( V_{BD} \) dependence. Usual way to compute required dependencies is based on numerical processing of integral relations (1) and (2) in each case. Impact ionization coefficients of electrons \( \alpha(E) \) and holes \( \beta(E) \) depend drastically on electric field \( E \). At the same time theoretical expressions for \( \alpha(E) \) and \( \beta(E) \) include usually some adjustable parameters. Therefore, to avoid large errors in calculating of multiplication factors, in computation of (1) and (2) are commonly used experimental dependences for \( \alpha(E) \) and \( \beta(E) \). Avalanche breakdown voltage \( V_{BD} \) is defined as applied
bias voltage at which multiplication factor of charge carriers tends to infinity (Sze, 1981), (Tsang, 1981), (Miller, 1955), (Grekhov & Serezhkin, 1980). Therefore, as seen from (2), breakdown condition is reduced to integral equation with \( m = 1 \) where field distribution \( E(x) \) is determined by solving Poisson equation. Bias voltage at which breakdown condition \( V = V_{BD} \) is satisfied can be calculated by method of successive approximations on computer. Thus, this method of determining \( V_{BD} \) and, hence, \( E(x) \) at \( V = V_{BD} \) requires time-consuming numerical calculations. The same applies to dependence \( M \) on \( V \). Similar calculations were performed for a number of semiconductor structures for certain thicknesses of diode’s base by which is meant high-resistivity side of \( p^+−n \) homojunction or narrow-gap region of heterojunction (Kim et al, 1981), (Stillman et al, 1983), (Vanyushin et al, 2007). In addition to great complexity, there are other drawbacks of this method of \( M(V) \) and \( V_{BD} \) determination – difficulties in application and lack of illustrative presentation of working results. Availability of analytical, more or less universal expressions would be very helpful to analyze different characteristics of devices with avalanche multiplication of charge carriers, for example, expression describing \( E(x) \), when we estimate tunnel currents in AHPDs. In this section are presented required analytical dependences (Osipov & Kholodnov, 1987), (Kholodnov, 1988), (Kholodnov, 1996-3). For quick estimate of breakdown voltage in abrupt \( p^+−n \) homojunction or heterojunction is often used well-known Sze-Gibbons approximate expression (Sze, 1981), (Sze & Gibbons, 1966):

\[
V_{BD} = A_V \times N^{-(s-2)/s}, \ V_s
\]

where

\[
s = 8, \ A_V = 6 \times 10^{13} \times \left(\frac{E_g}{1.1}\right)^{3/2}
\]

Gap \( E_g \) of semiconductor material forming diode’s base and dopant concentration \( N \) in it are measured in eV and cm\(^{-3}\), properly. As follows from Poisson equation, voltage value given by (10) corresponds to value of electric field at metallurgical boundary \( (x=0, \ \text{Fig. 2}) \) of \( p^+−n \) junction:

\[
E(0) = E_{BD} = A \times N^{1/s},
\]

where at \( s = 8 \)

\[
A = \sqrt{\frac{1.2 \times q}{\varepsilon \varepsilon_0}} \times \left(\frac{E_g}{\varepsilon \varepsilon_0}\right)^{3/4} \times 10^{10}
\]
$\varepsilon_0$ and $\varepsilon$ – dielectric constant of vacuum and relative dielectric permittivity of base material; $q$ – electron charge. Unless otherwise stated, in formulas (12) and (13) and below in sections 3.1-3.3 is used SI system of measurement units.

Figure 2. Schematic drawing of diode based on $p^+–n–n^+$ heterostructure (a) and distribution of electric field in it at avalanche breakdown voltage (b): 1 – $N_1 = N_1(0)$, 2 – $N_1 > N_1(0)$; $N_1$ – dopant concentration $n_{wg}$ in wide-gap layer I

Formulas (10) and (11) cannot be used for reliable estimates of $V_{BD}$ and $E_{BD}$ in semiconductor structures with thin enough base. Indeed, dependence of $V_{BD}$ on $N$ is due to two factors.
First, as follows from Poisson equation, the larger \( N \) the steeper the field \( E(x) \) decreases into the depth from \( x=0 \) comparing to value \( E_1=E(0) \) (Fig. 1b). Second, value of electric field \( E_1=E(0) \) at \( V=V_{BD} \) falls with decreasing of \( N \) due to decreasing of \( |\nabla E| \) in SCR. Drop of \( E(x) \) becomes more weaker with decreasing of \( N \) (Fig. 1b), therefore, at preset base’s thickness \( W \), initiation of avalanche process will require fewer and fewer field intensity \( E_1 \). At sufficiently low concentration \( N \), the lower the thicker \( W \) will be, variation of electric field \( E(x) \) on the length of base is so insignificant that probability of impact ionization becomes practically the same in any point of base. It means that breakdown voltage \( V_{BD} \) and field \( E_{BD} \) are independent on \( N \) and at the same time are dependent on \( W \), moreover, the thinner \( W \) then, evidently, the higher \( E_{BD} \). So using of formulas (10) and (11) at any values of \( W \), that done in many publications, contradicts with above conclusion. In next section 3.2 will be shown that value of breakdown field of stepwise \( \text{p}^+ - \text{n} \) junction in a number of semiconductor structures can be estimated by following formula:

\[
E_{BD}(N,W) = E_{BD}(0,W) \times \left[ 1 + \frac{N}{\bar{N}(W)} \right]^{1/s},
\]

where

\[
E_{BD}(0,W) = A \times \left( \frac{\varepsilon \varepsilon_0 A}{s \times q \times W} \right)^{1/(s-1)},
\]

\[
\bar{N}(W) = \left( \frac{\varepsilon \varepsilon_0 A}{s \times q \times W} \right)^{s/(s-1)}.
\]

It seen from expression (14) that at \( N < \bar{N}(W) \) electric field of avalanche breakdown \( E_{BD} \) is practically independent on dopant concentration \( N \) in diode’s base.

### 3.2. Avalanche breakdown field

Consider \( \text{p}^+ - \text{n}_w - \text{n}_g - \text{n}_w^+ \) heterostructure (Fig. 2). Symbols \( n_w \) and \( n_g \) indicate to unequal, in general, doping of high-resistivity layers of structure. Denote as \( W_1 \), \( W_2 \) and \( N_1 \), \( N_2 \) thicknesses of \( n_w \) and \( n_g \) layers and dopant concentrations in them, properly. Case \( W_2=0 \) corresponds to diode formed on homogeneous \( \text{p}^+ - \text{n} - \text{n}^+ \) structure. Let values \( N_1 \) and \( W_1 \) such that upon applying avalanche breakdown voltage \( V_{BD} \) to structure, SCR penetrates into narrow-gap \( n_g \) layer (Fig. 2). When \( W_1 \) and \( N_1 \) are small enough and \( W_2 \) is thick enough then avalanche process develops in \( n_g \) layer. In other words, with increasing bias
$V_B$ applied to heterostructure, electric field $E = E_2$ in narrow-gap layer on $n_{wg}/n_{ng}$ heterojunction (Fig. 2) reaches avalanche breakdown field $E_{2BD}$ in this layer earlier than electric field $E_1$ on metallurgical boundary ($x = 0$) of $p^+-n$ junction becomes equal to breakdown field $E_{1BD}$ in wide-gap $n_{wg}$ layer. This is due to the fact that at small values of $W_1$ and $N_1$ variation of field $E(x)$ within wide-gap layer is insignificant and probability of impact ionization in narrow-gap layer is much higher than in wide-gap. If, however, $W_1$ and $N_1$, $N_2$ are large enough and $W_2$ thin enough, then avalanche process is developed in wide-gap $n_{wg}$ layer only. For these values of thicknesses and concentrations electric field $E_1$ reaches value $E_{1BD}$ earlier than $E_2 - $ value $E_{2BD}$. Because of significant decreasing of electric field $E(x)$ in $n_{wg}$ layer with increasing distance from $x = 0$, field $E_2$ remains smaller $E_{2BD}$ despite the fact that band-gap $E_g$ in $n_{wg}$ layer is wider than band-gap $E_g$ in $n_{ng}$ layer. Distribution of electric field $E(x)$ in $n_{wg}$ and $n_{ng}$ layers of considered heterostructure is obtained by solving Poisson equation as defined by (6)-(9). When avalanche breakdown voltage $V_{BD}$ is applied to structure, then either $E_1 = E_{1BD}(N_1, W_1)$ or $E_2 = E_{2BD}(N_2, W_2)$. In section 3.1 is noted that at low enough concentrations $N_i$ avalanche breakdown fields $E_{iBD}(N_i, W_i)$ should not depend on $N_i$ and have definite value depending on $W_i$, where $i = 1, 2$. To account for this effect, formula (12) should be modified so that when $N \to 0$ then breakdown field $E_{BD}$ tends to some non-zero value. It would seem that it is enough to add some independent on $N$ constant to right side of expression (12). It is easy to see that such modification of formula (12) leads to contradiction. To verify that let’s consider situation when avalanche multiplication of charge carriers occurs in $n_{wg}$ layer, i.e. $E_1$ is close to $E_{1BD}$ and multiplication factor of holes $M_p$ (1) is fixed. Then, with increasing concentration $N_1$, field $E_1(W_1)$ (Fig. 2b) shall be monotonically falling function of $N_1$. Indeed, with increasing $N_1$, field $E_{1BD}$ and $|\nabla E_1(x)|$ are increasing also. Increasing $|\nabla E_1|$ must be such that when $x$ became larger some value $\bar{x}$ then value $E_1(\bar{x})$ has decreased (Fig. 2b). Otherwise, field $E(x)$ would increase throughout SCR that reasonably would lead to growth of $M_p$. This is evident from (1) and (2). On the other hand, adding constant to right side of expression (12) does not change $\partial E_{1BD}/\partial N_1$ and therefore results in, as follows from (6) and (9), non-monotonic dependence $E_i(W_i)$ on $N_i$. Equation (14) which can be rewritten for each of $n_{wg}$ and $n_{ng}$ layers as:

$$E_{iBD}(N_i, W_i) = A_i \times [N_i + \bar{N}_i(W_i)]^{1/s}$$

(17)

does not lead to that and other contradictions, From (17) follows that:

$$\bar{N}_i(W_i) = \left[ \frac{E_{iBD}(0, W_i)}{A_i} \right]^s$$

(18)
To determine dependences $E_{BD}(0, W)$, let’s consider behavior of $E_i(W)$ when parameters of heterostructure $N_1, N_2$ and $W_2$ are varying. From (6)-(9), (17) and (18) we find that when value

$$\Delta = N_2 + \bar{N}_2(W_2) - \left(\frac{e_1 \times A_1}{e_2 \times A_2}\right) \times \bar{N}_1(W_1) > 0$$

(19)

then avalanche breakdown is controlled by $n_{wg}$ layer. It means that

$$E_i(W_1) = E_{1BD}(N_1, W_1) - \frac{q N_1 \times W_1}{\varepsilon_0 \varepsilon_1}$$

(20)

If, however, $\Delta < 0$ then avalanche breakdown is controlled by $n_{wg}/n_{ng}$ heterojunction, i.e.

$$E_i(W_1) = \frac{\varepsilon_2}{\varepsilon_1} E_{2BD}(N_2, W_2)$$

(21)

From (17)-(21) we obtain that

$$\frac{\partial E_i(W_1)}{\partial N_1} \bigg|_{N_1 \to 0} = \begin{cases} \frac{A_2 \times \bar{N}_1^{(1/\beta)-1}}{s} \times \frac{q \times W_1}{\varepsilon_0 \varepsilon_1}, & \text{at } \Delta > 0 \\ 0, & \text{at } \Delta < 0 \end{cases}$$

(22)

Formulas (15) and (16) follow from expressions (18), (19) and requirement (23)

$$\lim_{\Delta \to 0} \left\{ \frac{\partial E_i(W_1)}{\partial N_1} \bigg|_{N_1 \to 0} \right\} = \lim_{\Delta \to 0} \left\{ \frac{\partial E_i(W_1)}{\partial N_1} \bigg|_{N_1 \to 0} \right\}$$

(23)

which means smoothness of field dependence $E(x)$ in real heterostructures, where parameters are varying continuously. Particularly, in semiconductors for which relations (11) and (13) are valid, breakdown field at metallurgical boundary of $p^{+}-n$ junction (or at heterojunction boundary, in narrow-gap layer of heterojunction, including isotype) can be described by formula

$$E_{BD}(N, W) = E_{BD}(0, W) \left[ 1 + \frac{N}{\bar{N}(W)} \right]^{-1/8}$$

(24)
where

$$E_{BD}(0, W) = X_e \times X_s^{-6/7} \times E_{BD}^{(InP)}(0, W); \quad \bar{N}(W) = X_e \times X_s^{-6/7} \times \bar{N}_{InP}(W)$$  \hspace{1cm} (25)

And values for InP semiconductor widely used for manufacturing of AHPDs (Tsang, 1981), (Stillman, 1981), (Filachev et al, 2010), (Filachev et al, 2011) are as follows:

$$E_{BD}^{(InP)}(0, W) = 4.3 \times 10^5 \times W^{-1/7} \text{ V/cm}; \quad \bar{N}_{InP}(W) = 3.4 \times 10^{15} \times W^{-6/7} \text{ cm}^{-3}$$  \hspace{1cm} (26)

$$X_e = 12.4 / e, \quad X_s = 1.35 / E_g$$ and gap $E_g$ in diode’s base is measured in eV and its thickness $W$ – in $\mu$m, respectively.

### 3.3. Avalanche breakdown voltage

It follows from expressions (6)-(9) and (14)-(16) that breakdown voltage $V_{BD}$ for $p^+ - n - n^+$ structure is given by expressions

$$V_{BD} = \frac{e \varepsilon_0}{2q} A^2 \times \left[ 1 + \frac{\bar{N}(W)}{N} \right]^{2/s} \times N \left( \frac{s-2}{s} \right) = A_V \times \left[ 1 + \frac{\bar{N}(W)}{N} \right]^{2/s} \times N \left( \frac{s-2}{s} \right), \text{ if } \frac{\bar{N}}{N} < \frac{1}{6}$$  \hspace{1cm} (27)

i.e. when diode’s base is not punch-through and

$$V_{BD}(N,W) = V_{BD}(0,W) \times \left[ \frac{N - \bar{N}(W)}{2s \times \bar{N}(W)} \right], \text{ if } \frac{\bar{N}}{N} > \frac{1}{6}$$  \hspace{1cm} (28)

i.e. when diode base is punch-through. In expression (28)

$$V_{BD}(0,W) = A \times \left( \frac{e \varepsilon_0 \times A}{s \times q} \right) \frac{1}{s-1} \times W^{s-1}$$  \hspace{1cm} (29)

Value of parameter $\theta$ is defined from equation $\theta = s \times (1 + \theta)^{1/s}$ and with good degree of accuracy it equals to $s^{(s-1)}$. Because $\theta > > 1$, therefore expression (27) practically coincides with formula (10), i.e. $V_{BD}$ of diode with thick base is independent on its thickness $W$. For diodes with thin base formed on semiconductors with parameters satisfying relations (11) and (14), namely when
breakdown voltage of diode depends on \( W \) and \( N \) as follows

\[
V_{BD}(N, W) = V_{BD}(0, W) \times \left[ \left( 1 + X_e^{4/7} \times X_g^{6/7} \times W^{8/7} \right) \times \frac{N}{2.65 \times 10^{15}} \right]^{1/8} - X_e^{4/7} \times X_g^{6/7} \times W^{8/7} \times \frac{N}{4.24 \times 10^{18}}
\]

(31)

where

\[
V_{BD}(0, W) = 43.1 \times X_e^{3/7} \times X_g^{-6/7} \times W^{6/7}
\]

(32)

In expressions (30)-(32) \( X_e = 12.4 / \epsilon \), \( X_g = 1.35 / E_g \) and gap \( E_g \) in base, dopant concentration in it \( N \) and thickness \( W \) is measured in eV, cm\(^{-3}\) and \( \mu \)m, respectively.

Avalanche breakdown voltage of double heterostructure discussed in Section 4 (Fig. 1) depends on relations between fundamental parameters of materials of \( n_{wg} \) and \( n_{ng} \) layers, their thicknesses and doping, and is determined, as follows from (6)-(9) and (14)-(16), by different combinations (with slight modification) of expressions (27)-(29) for these layers of heterostructure.

3.4. About correlation between impact ionization coefficients of electrons and holes

One of main goals of many experimental and theoretical studies of impact ionization phenomenon in semiconductors is to determine impact ionization coefficients of electrons \( \alpha(E) \) and holes \( \beta(E) \) as functions of electric field \( E \) (Sze, 1981), (Tsang, 1985), (Grekhov & Serezhkin, 1980), (Stillman & Wolf, 1977), (Dmitriev et al, 1987). Parameters of some semiconductor devices, for example, APDs (Sze, 1981), (Filachev et al, 2011), (Artsis & Kholodnov, 1984), (Stillman & Wolf, 1977) depend significantly on ratio \( K(E) = \beta(E) / \alpha(E) \). Performance of APD can be calculated on computer if \( \alpha(E) \) and \( \beta(E) \) are known (Sze, 1981), (Tsang, 1985), (Filachev et al, 2011), (Grekhov & Serezhkin, 1980), (Stillman & Wolf, 1977), (Dmitriev et al, 1987). Dependences \( \alpha(E) \) and \( \beta(E) \) are known, with greater or lesser degree of accuracy, for a number of semiconductors (Sze, 1981), (Tsang, 1985), (Grekhov & Serezhkin, 1980), (Stillman & Wolf, 1977), (Dmitriev et al, 1987). However in works concerned determination of impact ionization coefficients the problem of interrelation between \( \alpha(E) \) and \( \beta(E) \) has never been put. Even so, laws of conservation of energy and quasi-momentum in the act of impact ionization are maintained mainly by electron-hole subsystem of semiconductor (Tsang, 1985), (Grekhov & Serezhkin, 1980), (Dmitriev et al, 1987). Therefore, there is a reason to hypothesize some correlation between \( \alpha(E) \) and \( \beta(E) \), although perhaps not quite unique, for
example, owing to big role of phonons in formation of distribution functions. It is shown in this section that for number of semiconductors the following approximate relation is satisfied (Kholodnov, 1988)

\[
Z(E,\alpha(E),\beta(E)) = 9 \times 10^2 \times \left(\frac{10^5}{E}\right)^7 \times \frac{\alpha(E) - \beta(E)}{\ln \left[\frac{\alpha(E)}{\beta(E)}\right]} = C(E) \times Z_0 \approx \frac{e^3}{E_g^2}, \tag{33}
\]

Where: \(\varepsilon\) – relative dielectric permittivity, and gap \(E_g\), electric field \(E\), \(\alpha\) and \(\beta\) are measured in eV, V/cm and 1/cm, properly.

To derive relation (33) let’s consider thin \(p^+ - n - n^+\) structure in which thickness of high-resistivity base layer \(W\) satisfies to inequality

\[
W < W_0 = \frac{A\varepsilon\varepsilon_0}{q^2} \times N^{\frac{1-s}{s}} \tag{34}
\]

where \(\varepsilon_0\) – dielectric constant of vacuum; \(\varepsilon\) – relative dielectric permittivity of base material; \(q\) – electron charge; \(s\) and \(A\) – constants defining dependence of electric field \(E_{BD}\) = \(A \times N^{1/s}\) at metallurgical boundary \((x=0)\) of abrupt \(p^+ - n\) junction on dopant concentration \(N\) in base for avalanche breakdown in thick \(p^+ - n - n^+\) structure (Sections 3.1-3.3, (Sze, 1981), (Grekhov & Serezhkin, 1980), (Sze & Gibbons, 1966)). When condition (34) is satisfied then avalanche breakdown field can be written as

\[
E_{BD}(W) \approx A \times \left(\frac{A\varepsilon\varepsilon_0}{sqW}\right)^{\frac{1}{s-1}} \tag{35}
\]

And, under these conditions, variation of electric field \(E(x)\) along length of base \(W\) is so insignificant that probability of impact ionization is practically the same in any point of base of considered structure. For many semiconductors including Ge, Si, GaAs, InP, GaP relations given below are valid (Sze, 1981), (Kholodnov, 1988-2), (Kholodnov, 1996), (Sze & Gibbons, 1966)

\[
s = 8, A = \sqrt[\frac{1.2}{\varepsilon\varepsilon_0}] \times \left(\frac{E_g}{11q}\right)^{\frac{3}{4}} \times 10^{30}, \tag{36}
\]

In this case as it follows from (34) and (35)
And avalanche breakdown electric field for thin \(p^+ - n - n^+\) structure is defined by approximate universal formula

\[
E_{BD}(W) \approx \left( \frac{E_0^2}{e} \right)^{3/7} \times \frac{10^6}{W^{3/7}},
\]

(38)

In expressions (37) and (38) and below in this Section 3.4 concentration is measured in \(\text{cm}^{-3}\), energy – in eV, length – in μm, electric field – in V/cm. On the other hand condition of avalanche breakdown of \(p^+ - n - n^+\) structure (Sections 2, 3.1 and (Sze, 1981), (Tsang, 1985), (Grekhov & Serezhkin, 1980), (Stillman & Wolf, 1977))

\[
m(0, W) = \int_0^W \alpha(E(x)) \times \exp\left\{ \int_0^x \beta(E(x')) - \alpha(E(x')) dx' \right\} dx = 1,
\]

(39)

takes the form

\[
W \times [\alpha(E_{BD}) - \beta(E_{BD})] = \ln \left[ \frac{\alpha(E_{BD})}{\beta(E_{BD})} \right],
\]

(40)

That means the same probability of impact ionization in any point of diode’s base. And relation (33) follows from expressions (38) and (40). Let’s estimate applicable electric field interval for this relation. Expression (38) will be valid when inequality (41) is satisfied both for electrons and for holes

\[
E_{BD}(W) \times W > \left( \frac{W}{\lambda_R} \times E_R + E_{ion} \right) \times 10^4
\]

(41)

where \(\lambda_R\), \(E_{ion}\), \(E_R\) – mean free path for charge carriers scattered by optical phonons, threshold ionization energy of electrons or holes and energy of Raman phonon, respectively (Sze, 1981), (Tsang, 1985), (Grekhov & Serezhkin, 1980), (Stillman & Wolf, 1977), (Dmitriev, 1987). Taking into account that for many semiconductors
From (38) and (41) we find desired interval of electric field:

$$10^4 \times \frac{E_R}{\lambda_R} = E_{\text{BD}}(W_{\text{max}}) = E_{\text{min}} < E < E_{\text{max}} = E_{\text{BD}}(W_{\text{min}}) = 2 \times 10^6 \times \frac{E_g}{\sqrt{\varepsilon \times E_{\text{ion}}}}.$$  \tag{43}

Interval of electric field (43) is most often realized in experimental studies (Sze, 1981), (Tsang, 1985), (Grekhov & Serezhkin, 1980), (Stillman & Wolf, 1977), (Dmitriev, 1987). Ratio $W_{\text{min}}/\lambda_R$ is usually not more than a few units. Therefore, when $W < W_{\text{min}}$ then $E_{\text{BD}} = \frac{E_{\text{ion}}}{W} \times 10^4$ and hence when $E > E_{\text{max}}$ instead of (33) must be valid relation

$$\frac{E_{\text{ion}}}{E} \times \ln \left[ \frac{\alpha(E) - \beta(E)}{\alpha(E)} \right] = c(E) \approx 1.$$  \tag{44}

where $E_{\text{ion}}$ to be understood by largest in value threshold ionization energy of electrons and holes. On basis of relations (33) (or its upgraded version, if parameters $s$ and $A$ differ from values of (36)) and (44) can be obtained although approximate but relatively simple and universal analytical dependences of charge carriers multiplication factors and excess noise factors (Tsang, 1985), (Stillman et al, 1983), (Artsis & Kholodnov, 1984), (Woul, 1980), (McIntyre 1966), (Stillman & Wolf, 1977) on voltage as well as analytical expressions for avalanche breakdown voltage at different spatial distributions of dopant concentration in $p-n$ structures.

3.5. Miller’s relation for multiplication factors of charge carriers in $p-n$ structures

Usual way to calculate dependences of avalanche multiplication factors of charge carriers $M$ (Section 2) in $p-n$ structures on applied voltage $V_b$ is based on numerical processing of integral relations (1) and (2) in each case. Distribution of specific rate of charge carriers’ generation $g(x)$ in space charge region (SCR), i.e. when $-L_p < x < L_n$ (see inset in Fig. 3), is accepted in this Section 3.5 as exponential (and as special case – uniform). It is valuable for practical applications to have analytical, more or less universal, dependences $M$ on $V_b$. In article (Sze & Gibbons, 1966) was proposed analytical expressions for avalanche breakdown voltage $V_{\text{BD}}$, i.e. applied voltage value at which $M = \infty$, in asymmetric abrupt and linear $p-n$ junctions. Expression for $V_{\text{BD}}$ (Sze & Gibbons, 1966) in the case of asymmetric abrupt $p^+ - n$ junction was generalized in (Osipov & Kholodnov, 1987) for the case of thin $p^+ - n(p) - n^+$
structure (like as \( p-i-n \)) as discussed in Section 3.3. Using as model abrupt (stepwise) \( p-n \) junction under assumption that \( K(E)=\beta/\alpha = \text{const} \) (Kholodnov, 1988-2) has been shown that from (1), (2) and approximate relation (33), which is valid for number of semiconductors including Ge, Si, GaAs, InP, GaP, can be obtained analytical dependences of multiplication factors of charge carriers on voltage.

Figure 3: Dependences of exponents in Miller’s relation for electron \( n_e \) and holes \( n_h \) for “thick” abrupt \( p-n \) junction on applied voltage \( V \) at different values \( K = \beta/\alpha \) equal to 1, 2, 3, and 4.
Rewrite (33) in the form

\[ \alpha(E) \frac{K(E) - 1}{\ln K(E)} = \frac{5}{6} \times \left( \frac{\varepsilon \varepsilon_0}{6 \times 10^8 \times q} \right)^3 \times \left( \frac{11 \varepsilon_0}{E_g} \right)^\frac{1}{6} \times \left( \frac{E}{10^5} \right)^7 \]  

(45)

In (45) and below in this Section 3.5 is accepted (unless otherwise specified) the following, convenient for this study, system of symbols and units (Sze, 1981): gap \( E_g \) and threshold ionization energy \( E_{ion} \) in eV; electric field \( E \) in V/cm; bias \( V_b \) in V; multiplication factors \( \alpha \) and \( \beta \) in cm\(^{-1}\); electron charge \( q \) in C; dielectric constant of vacuum \( \varepsilon_0 \) in F/m; concentration including shallow donors \( N_D \) and acceptors \( N_A \) in cm\(^{-3}\); concentration gradient \( a \) in cm\(^{-4}\); width of SCR \( L_p \) and \( L_n \) in \( p \) and \( n \) layers and thicknesses of these layers (inset in Figure 3) in \( \mu \)m, light absorption coefficient \( \gamma \) in cm\(^{-1}\). In this section, analytical dependences \( M(V) \) in \( p-n \) structures have been calculated under no \( K(E) = \text{const} \) condition. Such calculations are possible because ratio \( \frac{K(E) - 1}{\ln K(E)} \) varies, typically, much slighter than \( E \). In some cases it allows using relation (45) to integrate analytically (in some cases – approximately) expressions (1) and (2) and, thus, get analytical, more or less universal, relatively simple dependences \( M(V) \). The most typical cases are considered: abrupt (stepwise) and gradual (linear) \( p-n \) junctions like as in model given in (Sze, 1987), (Sze & Gibbons, 1966) and thin \( p-p^-n(p^-n)^+ \) structure (like as \( p-i-n \)) with stepwise doping profile as in model presented in (Osipov & Kholodnov, 1987). For purposes of discussion and comparison of obtained results with numerical calculations and experimental data, multiplication factors will be written in traditional common form

\[ M_n = \frac{1}{1 - v^{n_p}}, M_p = \frac{1}{1 - v^{n_n}}, M = \frac{1}{1 - v^{n_N}}, \]  

(46)

where \( v = V / V_{BD} \). This form was first proposed by Miller in 1955 (Miller, 1955) and then, despite lack of analytical expressions for exponents \( n_p, n_n, \hat{n} \), has been widely used as “Miller’s relation” (Sze, 1981), (Tsang, 1985), (Grekhov & Serezhkin, 1980), (Leguerre & Urgell, 1976), (Bogdanov et al, 1986). It was found that values of these exponents depend on many factors including, in general, voltage as well (Kholodnov, 1988-2), (Grekhov & Serezhkin, 1980), Fig. 3. Form of writing (46) clearly shows that \( M(V) \rightarrow \infty \) when \( V \rightarrow V_{BD} \).

3.5.1. Stepwise \( p-n \) junction

In this case from relations (1), (2) and (45) and Poisson equation (SI units)

\[ \frac{dE}{dx} = \begin{cases} \frac{q N_A \varepsilon_0}{\varepsilon E_0}, & x < 0 \\ \frac{q N_D \varepsilon_0}{\varepsilon E_0}, & x > 0 \end{cases} \]  

(47)
follow that

\[ M_n = \frac{(K_0 - 1)}{(K_0 - K_0^*)}, \quad M_p = k_0^* \times M_n \]  

(48)

\[ V_{BD} = 6 \times 10^{13} \times \left( \frac{E_g}{1} \right)^{3/2} \times N_{eff}^{-3/4}, \quad N_{eff} = \frac{N_A \times N_D}{N_A + N_D} \]  

(49)

where \( K_0 \) – value \( K(x) \) when \( E(x) = E(0) = E_0 \), i.e. value of \( K \) at metallurgical boundary of \( p-n \) junction (see inset in Fig. 3). Formula (49) for \( V_{BD} \) at \( N_D \ll N_A \) or \( N_A \ll N_D \) becomes well-known Sze-Gibbons relation (Sze, 1981), (Sze & Gibbons, 1966). If charge carriers are generated uniformly in SCR then computations lead to following expressions:

\[ \frac{\hat{M}}{M_n} = N_A \times \exp[\xi_A] \times \left[ e(K_0 - 1) + \frac{\xi_D}{\xi}_D \right] \times \exp[\xi_D] \times e(1 - K_0) + \frac{\xi_A}{\xi}_A, \]  

where

\[ \xi_A,D(\nu) = (N_{eff} / N_{eff}) \times \nu^4 \times \ln K_0 \leq 1; \]  

(50)

\[ \hat{M} = \left( 1 - K_{eff} \right)^{-1} \times M_n, \]  

(52)

when

\[ e(x) = \text{unity function} \quad (Zeldovich & Myshkis, 1972), \quad K_{eff} = K_0 + K_0^{-1}. \]  

Expression (50) is obtained by expanding the function \( Y(x, -L_p) \) as a power series in

\[ \int_{-L_p}^{x} (\beta - \alpha) dx' \]

and expression (52) was derived by standard method of integrating fast-changing functions (Zeldovich & Myshkis, 1972).
3.5.2. Gradual (linear) $p-n$ junction

In this case Poisson equation can be written as (SI units):

$$\frac{dE}{dx} = \frac{q \cdot \sigma \times x}{\varepsilon \varepsilon_0} \times x \quad (54)$$

where $\sigma$ - slope of linear concentration profile

and therefore

$$M_n = (K_0 - 1) / (K_0 - K_0^{3/5}), M_p = K_0^{3/4} \times M_n \quad (55)$$

$$V_{BD} = 60 \times \left(\frac{3 \times 10^{20}}{\sigma}\right)^{2/5} \times \left(\frac{E_x}{1.1}\right)^{6/5} \times \left(\frac{17.7}{\varepsilon}\right)^{1/5} \quad (56)$$

In derivation of relations (55) and (56) was used known expression for voltage distribution on linear $p-n$ junction (Sze, 1981) and was also taken into account that (Gradstein & Ryzhyk, 1963)

$$\int_0^y \frac{(y-x)^7}{\sqrt{x}} dx = \frac{4096}{6435} y^{15/2} \quad (57)$$

Formula (56) differs from known formula Sze-Gibbons for avalanche breakdown voltage of linear $p-n$ junction (Sze, 1981), (Sze & Gibbons, 1966) by last multiplicand, which for typical values of $\varepsilon \approx 10$ (Sze, 1981), (Casey & Panish, 1978) is close to unity.

3.5.3. Thin $p^+ - n(p) - n^+$ structure ($p-i-n$)

When thickness of high-resistivity region (base) of considered structure

$$W > \bar{W} = \sqrt[3/4]{\frac{E_x}{1.1} \times \frac{10^{10}}{N^{7/8}} = 2 \sqrt{\varepsilon} \times E_x^{3/4} \times \left(\frac{3 \times 10^{15}}{N}\right)^{7/8}} \quad (58)$$

where $N$ – dopant concentration (for example, donor) in base, and when $V_b = V_{BD}$ then SCR does not extend to entire thickness of base ((Osipov & Kholodnov, 1987), Sections 3.1-3.3, inset in Fig. 4). In this case, expressions (48)-(53) remain apparently valid. In opposite case, base is depleted by free charge carriers when $V_b < V_{BD}$ that gives in the result substantially
other expressions for avalanche multiplication factors of charge carriers and avalanche breakdown voltage. When \( W < \bar{W} \) then from relations (1), (2) and (45) and Poisson equation

\[
\frac{dE}{dx} = -\frac{qN}{\varepsilon \varepsilon_0}
\]

(59)

**Figure 4.** Dependences of analytical (solid lines) and numerical (dashed lines) (Leguerre & Urgell, 1976) limiting values of exponent \( n_B = \lim_{V \to V_{BD}} n(V) \) in Miller’s relation (46) on concentration of donor dopant \( N_D \) in “thick” high-resistivity layer of stepwise \( p^+-n^-n^+ \) structure, 1 – Si, 2 – Ge, 3 – GaAs, 4 – GaP. Values \( K(E) \), as in (Leguerre & Urgell, 1976), are taken from (Sze & Gibbons, 1966). In inset – scheme of “thick” \( p^+-n^-n^+ \) structure.
we find that

\[ M_n = (K_0 - 1) / (K_0 - K_0^{p^+}), M_p = K_0^{p^+} \times M_n, \quad (60) \]

where

\[ \delta^8 = \frac{(V + V_t)^8 - (V - V_t)^8}{V_2^8}, \quad (61) \]

\[ V_1 = \frac{qNW^2}{2\varepsilon\varepsilon_0} \times 10^{-6}, \quad (62) \]

\[ V_2 = \left( \frac{6\varepsilon\varepsilon_0}{5 \times 10^3 qW^2} \right)^4 \times \left( \frac{1.1}{E_g} \right)^6 \times \frac{1}{N}, \quad (63) \]

In deriving expressions (60)-(63), multiplication of charge carriers in \( p^+ \) and \( n^+ \) layers and voltage drop on them is considered negligible. This is justified because of significant decreasing of electric field \( E(x) \) deep into high-doped layers of the structure (Sze, 1981), (Kholodnov 1996-1), (Kholodnov 1998), (Leguerre & Urgell, 1976). Admissibility of such neglect is confirmed also by formula (49) when \( N_A < < N_D \) or when \( N_D < < N_A \). Avalanche breakdown voltage is determined by equation \( \tilde{\delta} = 1 \) which has no exact analytical solution. However, till \( W \) surpasses \( \tilde{W} / 8 \), then value of field at \( x = W \) is much less than value of field at \( x = 0 \). In this case, using smallness parameter

\[ \left( 1 - 2 \times \frac{V_1}{V_2} \right)^8 \ll 1, \quad (64) \]

we find that in zeroth-order approximation with respect to this parameter

\[ V_{BD} = V_2 - V_1. \quad (65) \]

In the case of very thin base when

\[ W \leq W_0 = \frac{1}{8} \tilde{W}, \quad (66) \]
electric field varies so slightly along base that probability of impact ionization is practically the same in any point of it ([Osipov & Kholodnov, 1987], [Kholodnov 1988-1], Sections 3.2 and 3.4). As a result

\[ M_n = (K - 1) / (K - K^n), \quad M_p = K^n \times M_n. \]  

(67)

\[ M = \frac{\gamma W}{\gamma W + v^7} \times \frac{K^n \times \exp(\gamma W) - 1}{\exp(\gamma W) - 1} \times M_n. \]  

(68)

\[ V_{BD} = 7 \times \left( \frac{3}{25} \times \left( \frac{3q}{50\pi e_0} \right)^{3/2} \left( \frac{E_g}{W} \right)^{6/7} \times 10^6 \right) = 98 \times \left( \frac{W \times E_g}{\sqrt{\gamma}} \right)^{6/7}, \]  

(69)

where \( \gamma < 0 \) if structure is illuminated through \( p^+ \) region (front-side illuminated) and \( \gamma > 0 \) if structure is illuminated through \( n^+ \) region (back-side illuminated).

3.6. Discussion of the results. Comparison with computed and experimental data

3.6.1. To formulas for avalanche breakdown electric field and voltage for abrupt \( p^- n \) junction

In sections 3.1-3.3 were derived approximate universal formulas for avalanche breakdown field \( E_{BD} \) and voltage \( V_{BD} \) for abrupt \( p^- n \) junction taking into account finite thickness of high-resistivity layer \( W \). Comparative values of breakdown field \( E_{BD}(0, W) \) for Si, Ge and InP most often used for fabrication of APDs computed by formulas (25) and (26) and found from numerical solution of breakdown integral equation \( m = 1 \), where \( m \) is defined by (2) are shown on Fig. 5 (Sze, 1981), (Tsang, 1985), (Stillman, 1981), (Filachev et al, 2010), (Filachev et al, 2011), (Groves et al, 2005), (Stillman et al, 1983), (Trommer, 1984), (Woul, 1980), (Leguerre & Urgell, 1976), (Bogdanov et al, 1986), (Gasanov et al, 1988), (Brain, 1981), (Tager & Vald-Perlov, 1968). It is seen that in the most practically interesting range \( W = (0.2 \div 10) \mu m \) for all a.m. semiconductors analytical \( E_{BD}(0, W) \) and calculated \( E_{BD}(W) \) values of breakdown field differ by less than 20 %. Relatively drastic fall of ratio \( E_{BD}(0, W)/E_{BD}(W) \) in comparison to unity with decrease of \( W \) (for thin enough \( W \)) is due to the fact that, as shown in Sec. 3.4, if

\[ W < W_{min} = 5 \times 10^{-3} \times \frac{\sqrt{\gamma}}{E_g} \times E_{un}^{7/6}, \]  

(70)

then formulas (25) and (26) are not true. To estimate breakdown field \( E_{BD}(0, W) \) at values \( W \) defined by (70) can be used the following formula
Figure 5. Dependence of ratio between analytical value of breakdown field $E_{BD}(0, W)$ obtained by formulas (25) and (26) and numerical value $E^{(0)}_{BD}(0, W)$. Dashed curve – analytical value of effective avalanche breakdown field $E_{BD}^*(0, W) = X_{\epsilon}^{3/7} \times X_{g}^{6/7} \times E_{BD}(0, W)$. Curves 1 and 1’ – Si, 2 – InP, 3 – Ge. Values $\alpha(E)$ and $\beta(E)$ are taken: for curves 1 and 3 – from Table 1 of monograph (Grekhov & Serezhkin, 1980), for curve 1’ – from (Kuzmin et al, 1975), for curve 2 – from (Cook et al, 1982).

If assume that in Si threshold energy of impact ionization $E_{\text{th}}$ of holes is higher than electrons, and it equals to 5 eV (Sze, 1981), then from (70) we find for Si $W_{\text{min}}=0.1 \, \mu m$. Estimates based on data from studies (Sze, 1981), (Tsang, 1985), (Stillman et al, 1983), (Grekhov & Serezhkin, 1980), (Stillman & Wolf, 1977) show that for Ge and InP value $W_{\text{min}}$ is 2-3 times smaller.

Therefore curve 1 in Fig. 5 starts to fall significantly below unity at larger values $W$ than curves 2 and 3. Analytical and computed dependences $E_{BD}$ on $N$ for InP used in high-performance APDs for wavelength range $\lambda=(1+1, 7) \, \mu m$ as wide-gap layers in double hetero-
structures (Fig. 1, 2) are shown on Fig. 6 (Tsang, 1985), (Stillman, 1981), (Filachev et al, 2010), (Forrest et al, 1983), (Filachev et al, 2011), (Stillman et al, 1983), (Ando et al, 1980), (Trommer, 1984). It is seen that $E_{BD}^{(a)}(N, W)$ and $E_{BD}^{(c)}(N, W)$ differ from each other by less than 10 %. In Fig. 7 and 8 are shown universal dependences of breakdown voltage $V_{BD}^{(a)}$ on $N$ and $W$ calculated by formulas (11), (27)-(29). It is seen from Fig. 7 that Sze-Gibbons relations (10) (Sze, 1981), (Sze & Gibbons, 1966) can be used to determine $V_{BD}$ when $N > N_{min} = 10 \times \bar{N}(W)$ only. Value of this minimal concentration, for example, for classic semiconductors Si, Ge, GaAs, GaP and InP at $W = (1-2) \mu$m equals to $(1\pm 5) \times 10^{16}$cm$^{-3}$. As shown on lower inset in Fig. 7, dependence $V_{BD}$ on $N$ is in the strict sense non-monotonic. Such kind of dependence $V_{BD}$ on $N$ is due to the fact that for small enough $N$ breakdown field $E_{BD}$ is growing faster with increasing $N$ than $|\nabla E| \propto N$ in diode’s base. Maximum $V_{BD}$ is reached, as it follows from (28), at

$$N = N_{max} = \left( \frac{s}{2^{s-1}} - 1 \right) \times \bar{N}(W)$$  \hspace{1cm} (72)

and expressed as

$$V_{BD}^{(max)} = \left( s-1 \right) \times 2^{s-1} + 1 \times (2s)^{-1} \times V_{BD}(0, W)$$  \hspace{1cm} (73)

when $s=8$, value $N_{max} = 1.2 \times \bar{N}$, $\Delta V_{max} = 2.86 \times 10^{-2} < 1$ and absolute value $\Delta V_{max}$ can reach tens Volts, and even more (see Fig. 7). The analytical dependences $V_{BD}^{(a)}(N, W)$ (Fig. 7 and 8) for a number of semiconductors are in good agreement with $V_{BD}^{(c)}(N, W)$ computed on the basis of integral equations (1) and (2) (Sze, 1981), (Tsang, 1985), (Stillman, 1981), (Stillman et al, 1983), (Grekhov & Serezhkin, 1980), (Leguerre & Urgell, 1976). Note that results of comparison $V_{BD}^{(a)}(N, W)$ with $V_{BD}^{(c)}(N, W)$ and $E_{BD}^{(a)}(N, W)$ with $E_{BD}^{(c)}(N, W)$ depend on accuracy of determination of impact ionization coefficients of electrons $\alpha(E)$ and holes $\beta(E)$ which are sharp functions of electric field $E$. As a rule, different authors obtain different results (Sze, 1981), (Tsang, 1985), (Stillman, 1981), (Stillman et al, 1983), (Grekhov & Serezhkin, 1980), (Sze & Gibbons, 1966), (Stillman & Wolf 1977), (Dmitriev et al, 1987), (Tager & Vald-Perlov, 1968), (McIntyre, 1972), (Cook et al, 1982) (see, for example, curves 1 and 1’ in Fig. 5). In addition, deducing of relations (1) and (2) is based on local relation between $\alpha$ and $\beta$ (Sze, 1981), (Tsang, 1985), (Stillman, 1981), (Filachev et al, 2011), (Stillman et al, 1983), (Grekhov & Serezhkin, 1980), (Sze & Gibbons, 1966), (Stillman & Wolf 1977), (Dmitriev et al, 1987), (Tager & Vald-Perlov, 1968), (McIntyre, 1972), (Cook et al, 1982) which is not always valid (McIntyre, 1972), (Gribnikov et al, 1981), (Okuto & Crowell, 1974), (McIntyre, 1999).
Figure 6. Dependence of field $E_{BD}$ on $N$ for InP: 1 – $W=0.5 \, \mu m$, 2 – $W=2 \, \mu m$, 3 – $W=8 \, \mu m$. Solid lines – formulas (24)-(26), dashed curves – numerical calculation. Values $\alpha(E)$ and $\beta(E)$ are taken from (Cook et al, 1982). In inset is shown dependence of effective concentration $\tilde{N} = X_{\epsilon}^{4/7} \times X_{\sigma}^{6/7} \times \tilde{N} \equiv \tilde{N}_{\text{InP}}$ on $W$. Concentration is measured in cm$^{-3}$, field – in V/cm and thickness $W$ – in $\mu$m.
Figure 7. Dependence of avalanche breakdown voltage $V_{BD}$ of homogeneous $p^+ - n - n^+$ structure on dopant concentration $N$ in base: solid line – (31) and (32), dotted line – expressions (10) and (11). In lower inset: dependence of relative voltage $\Delta V_{\text{rel}} = V_{BD} / V_{BD}(0, W) - 1$ normalized to concentration $\tilde{N}(W)$ at $N \leq 4 \times \tilde{N}(W)$. In upper inset: dependence of effective $\Delta V_{\text{max}}^* = X_e^{3/7} \times X_g^{6/7} \times [V_{BD} - V_{BD}(0, W)]_{\text{max}} = \Delta V_{\text{max}}^{\text{InP}}$ on base thickness $W$. Voltage is measured in V, thickness $W$ – in µm.
Figure 8. Dependence of effective avalanche breakdown voltage $V_{BD}^* = X^*_{\epsilon} \times X^*_{g} \times V_{BD} \equiv V_{BD}^{(0)}$ of homogeneous $\rho^+ - n - n^+ - n^+$ structure on thickness of its base $W$ for three values of effective concentration $N^* = X^*_{\epsilon} \times X^*_{g} \times N \equiv N_{\text{InP}}$:

1. $N^* = 3 \times 10^{16} \text{cm}^{-3}$, 2. $N^* = 3 \times 10^{14} \text{cm}^{-3}$, 3. $N^* = 3 \times 10^{12} \text{cm}^{-3}$. In inset is shown dependence $\tilde{W}$ on $N^*$. Concentration is measured in cm$^{-3}$, voltage $V$ in V, thickness $W$ in $\mu$m.
3.6.2. To correlation between values of impact ionization coefficients of electrons and holes

In Section 3.4 is shown that there is reason to suppose existence of some correlation between values of impact ionization coefficients of electrons $\alpha(E)$ and holes $\beta(E)$, and form of required relation (expression (33) and (45)) is proposed. It is obvious from Fig. 9 that values $Z_0 = \varepsilon_3 / E_g$ may differ by many orders of magnitude in different semiconductors. At the same time, for presented in Fig. 9 Ge, Si and GaP, function $c(E)$ (see relations (33) and (45)) in range of fields where $\alpha(E)$ and $\beta(E)$ vary in several orders of magnitude (Okuto & Crowell, 1975), remains, as it follows from (33) and (45), of the order of unity. Calculations based on experimental dependences $\alpha(E)$ and $\beta(E)$ (Cook et al, 1982) show that in $InP$ value $c(E)$ is some more closely to 1. It is evident from Fig. 10 that for GaAs, regardless of orientation of crystal with respect to electric field, function $c(E)$ depends weakly on $E$ in comparison with impact ionization coefficients of charge carriers (which values are taken from (Lee & Sze, 1980)), and differs from unity by no more than 2-3 times. A similar situation takes place in Ge (Fig. 11, according to (Mikawa et al, 1980)). As shown in (Kobajashi et al, 1969) dependences $\alpha(E)$ and $\beta(E)$ measured in (Miller, 1955), (McKay & McAfee, 1953) in the range of fields $E=(1.5\pm2.7)\times10^5$ V/cm can be described in Ge by formulas $\alpha(E)=7.81\times10^{-34}\times E^7$, $\beta(E)\cong 2\alpha(E)$. This result agrees well with expression (33). Note that, $c(E)$ differs from unity approximately by the same factor, as values $\alpha(E)$ and $\beta(E)$ for the same material obtained by different authors differ, respectively, from each other (Sze, 1981), (Tsang, 1985), (Forrest et al, 1983), (Grekhov & Serezhkin, 1980), (Sze & Gibbons, 1966), (Stillman & Wolf 1977), (Dmitrieve et al, 1987), (Tager & Vald-Perlov, 1968), (Cook et al, 1982), (Okuto & Crowell, 1974), (Okuto & Crowell, 1975), (Lee & Sze, 1980), (Mikawa et al, 1980), (Kuzmin et al, 1975). Using procedure described in Section 3.4, we can also determine relation between $\alpha(E)$ and $\beta(E)=K(E)\times \alpha(E)$ in the case when relations (11) and (13) are not satisfied (Grekhov & Serezhkin, 1980). It seems, relation required for such case, i.e. under assumption of power dependence $\alpha$ on $E$ and $K(E)=const$, was obtained for the first time in (Shotov, 1958).

3.6.3. To Miller’s relation

From (48), (55) and (67) follow that, exponents in Miller’s relation (46) for multiplication factors of electrons and holes are given by

$$n_n \times \ln v = \ln \left[ \frac{(K_0^\xi - 1)}{(K_0 - 1)} \right],$$

$$n_p \times \ln v = \ln \left[ \frac{K_0}{K_0 - 1} (1 - K_0^{-\xi}) \right],$$

where $\xi = 4, 5$ and 7 for stepwise $p-n$ junction, linear $p-n$ junction and very thin (66) $p^-n-n^+$ structure (situation 1, 2 and 3, respectively). If thickness of base in $p^-n-n^+$ structure is not very small, i.e., $W_0 < W < \tilde{W}$ (situation 4) then as it follows from formula (60), ex-
ponents \( n_n \) and \( n_p \) are also expressed by (74) and (75) but in right side of those expressions \( \bar{v} \) substitutes \( v \) and \( \xi = 8 \). Value of exponent \( \bar{n} \) lies between values \( n_n \) and \( n_p \). From (1) and (2) apparent that when \( \alpha = \beta \) then factors \( M_n, M_p \) and \( \bar{M} \) coincide with each other, i.e., \( n_n = n_p = \bar{n} = n \), and, as it follows from expressions (74) and (75), regardless of bias voltage applied, \( n = 4, 5 \) and 7 for situations 1, 2 and 3, respectively. Exponents in Miller’s relation have the same values when \( V << V_{BD} \) more exactly, when \(|\ln K_0/\ln v| << \xi\), regardless of ratio \( K_0 = \beta(\alpha)/\alpha(\beta) \). When \( V \to V_{BD} \) or more exactly, if

\[
\Delta v = 1 - v < \min\left[\frac{1}{\xi} - 1, \frac{1}{\xi}\right], \quad M > > 1
\]

Then for these situations

\[
n_n = n_{nB} = \xi \times K_0 \times \frac{\ln K_0}{(K_0 - 1)}, \quad n_p = n_{pB} = \xi \times \frac{\ln K_0}{(K_0 - 1)}.
\]

Figure 9. Dependence \( Z(E) \) [relation (33)] in Ge, Si, and GaP for \( \alpha(E) \) and \( \beta(E) \) from (Okuto & Crowell, 1975)
Graphs in Fig. 4 allow comparing numerical values of exponents $n_{B}$ and $n_{pB}$ calculated in (Leguerre & Urgell, 1976) and analytical $n_{B}^{(a)}$ computed by formulas (76) for asymmetrical stepwise $p-n$ junction. Like as in (Leguerre & Urgell, 1976), experimentally determined functional dependencies $\alpha(E_{0})$ and $\beta(E_{0})$ (Sze & Gibbons, 1966) were used in calculations of dependences $n_{B}^{(a)}$. As follows from (46), when $M > > 1$, then ratio of analytical value of multiplication factor $M_{(a)}$ to calculated $M_{(c)}$ equals to ratio $n_{B}^{(c)}$ to $n_{B}^{(a)}$ (Fig. 11-13). It obviously from Fig. 11-13 that for all considered semiconductors (with curves $\alpha(E)$ and $\beta(E)$ taken from (Sze & Gibbons, 1966)), dependences $M^{(B)}(V)$ and $M^{(C)}(V)$ do not differ by more than 50 %. Dependences of exponents $n_{B}^{(a)}$ and $n_{p}^{(a)}$ on voltage and $n_{B}^{(c)}$ and $n_{p}^{(c)}$ on ratio $K = \beta / \alpha$ are illustrated in Fig. 3 and 14, respectively. It should be noted that numerical values of exponent in Miller’s relation, as well as, value $V_{BD}$ depend, obviously, on what functions $\alpha(E)$ and $\beta(E)$ are used in (1) and (2) in calculations. Let’s take the simplest case when $\alpha(E) = \beta(E)$ and $p-n$ junction is stepwise. Varying expressions (1) and (2), we find that under considered conditions

$$n_{B} = \frac{\varepsilon E_{0}}{500 \times q \times N_{eff}} \times \alpha(E_{BD}) \times E_{BD},$$

(77)

where $E_{BD} = E(0)$ at $V = V_{BD}$ is determined from condition

$$\int_{0}^{E_{BD}} \alpha(E) dE = \frac{100}{\varepsilon E_{0}} \times N_{eff},$$

(78)

In Fig. 15a are shown dependences $n_{B}(N_{eff})$ calculated from relations (77) and (78) for four values $\alpha(E) = \beta(E)$ obtained for GaAs by different authors (Grekhov & Serezhkin, 1980), (Okuto & Crowell, 1975), (Kressel & Kupsky, 1966), (Nuttall & Nield, 1974). It is seen that analytical value $n_{B}^{(a)} = n_{pB}^{(a)} = 4$ calculated by formulas (76) approximately equals to mean value with respect to curves 1-4 in Fig. 15a. According to obtained above results expressions (48)-(53) are not valid when concentration $N_{eff} \geq (N_{eff})_{max} \cong 2 \times 10^{17} \times (E_{g})^{2} \times E_{ion}^{-4/3}$

(79)

which for many semiconductors is of the order of $10^{15}$ cm$^{-3}$. At such high concentrations, as it follows from Section 3.4 and (Kholodnov, 1988-1) and relations (1) and (2), for stepwise $p-n$ junction

$$n_{B} = \frac{\ln \left( K_{0}^{v} - 1 \right) / \left( K_{0} - 1 \right)}{\ln v}, n_{p} = n_{a} + \frac{1 - v}{\ln v} \ln K_{0},$$

(80)
Figure 10. Dependence $C(E)$ at different orientations of GaAs crystal with respect to electric field for values $\alpha(E)$ and $\beta(E)$ from (Lee & Sze, 1980)
Figure 11. Dependence of ratio between analytical values of avalanche multiplication factors $M^{(a)}$ of electrons and holes and numerical values $M^{(c)}$ (Leguerre & Urgell, 1976) in stepwise asymmetric Ge $p-n$ junction on value of multiplication factor $M=M^{(a)}$ of charge carriers. Solid lines – electrons, dashed – holes. Dopant concentration in high-resistivity part of $p-n$ junction $N$, cm$^{-3}$: 1 – $1 \times 10^{15}$, 2 – $3 \times 10^{15}$, 3 – $10^{16}$, 4 – $3 \times 10^{16}$, 5 – $6 \times 10^{16}$. Values $K(E)$, as in (Leguerre & Urgell, 1976), are taken from (Sze & Gibbons, 1966).
Figure 12. Dependence of ratio between analytical values of avalanche multiplication factors $M^{(a)}$ of electrons and holes and numerical values $M^{(c)}$ (Leguerre & Urgell, 1976) in stepwise asymmetric Si $p$–$n$ junction on value of multiplication factor $M = M^{(a)}$ of charge carriers. Solid lines – electrons, dashed – holes. Dopant concentration in high-resistivity part of $p$–$n$ junction $N$, cm$^{-3}$: 1 – $10^{15}$, 2 – $3 \times 10^{15}$, 3 – $10^{16}$, 4 – $3 \times 10^{16}$, 5 – $6 \times 10^{16}$. Values $K(E)$, as in (Leguerre & Urgell, 1976), are taken from (Sze & Gibbons, 1966)
Figure 13. Dependence of ratio between analytical values of avalanche multiplication factors $M^{(a)}$ of electrons and holes and numerical values $M^{(c)}$ (Leguerre & Urgell, 1976) in stepwise asymmetric GaAs (solid lines) and GaP (dashed lines) $p-n$ junctions on value of multiplication factor $M = M^{(a)}$ of charge carriers. Solid lines – electrons, dashed – holes. Dopant concentration in high-resistivity part of $p-n$ junction $N$, cm$^{-3}$: 1 – $10^{15}$, 2 – $3 \times 10^{15}$, 3 – $10^{16}$, 4 – $3 \times 10^{16}$, 5 – $6 \times 10^{16}$. Values $K(E)$, as in (Leguerre & Urgell, 1976), are taken from (Sze & Gibbons, 1966).
Figure 14. Dependence of limiting values $n_p = \lim_{V \rightarrow V_{BD}} n(V)$ of exponents in Miller’s relation for electron $n_n$ and holes $n_p$ for “thick” abrupt $p-n$ junction on $K = \beta / \alpha$. 
moreover

\[ n_n B = K_0 \ln \left( \frac{K_0}{(K_0 - 1)} \right) = K_0 \times n_p B. \]  (81)

\[ \text{Figure 15. Dependences } n_B(N_{\text{eff}}) \text{ in GaAs calculated on the base of different dependences } \alpha(E) = \beta(E), \text{ taken from: } 1 \text{ – (Shabde & Yeh, 1970), } 2 \text{ – (Grekhov & Serezhkin, 1980), } 3 \text{ – (Okuto & Crowell, 1975), } 4 \text{ – (Kressel & Kupsky, 1966), } 5 \text{ – (Sze & Gibbons, 1966). Dashed lines – analytical values.} \]

For comparison, in Fig. 15b are presented dependences of \( n_B(c) N_{\text{eff}} \) and \( n_B(a) N_{\text{eff}} \) for the case \( \alpha = \beta \), when \( n_n B = n_p B = n_B \). It is seen that value \( n_B(a) N_{\text{eff}} \) is approximately equal to mean value with reference to curves 2, 3 and 5 in Fig. 15b plotted on the base of numerical data. Note that starting from \( N_{\text{eff}} \cong (N_{\text{eff}})_{\text{max}} \), breakdown voltage \( V_{BD} \) dependence on \( N_{\text{eff}} \) becomes, with growth \( N_{\text{eff}} \), more and more weaker than that described by equation (49), and in limit tends to
value $V_{BD} = E_{im}/q$. This conclusion accords with results of studies (Grekhov & Serezhkin, 1980), (Nuttall & Nield, 1974). Obtained results agree well with experimental results for a number of $p$–$n$ structures, including based on $Ge$, $Si$, $GaAs$, $GaP$ (Sze, 1981), (Tsang, 1985), (Stillman et al, 1983), (Miller, 1955), (Grekhov & Serezhkin, 1980), (Sze & Gibbons, 1966), (Stillman & Wolf, 1977), (Bogdanov et al, 1986), (Cook et al, 1982), (Shotov, 1958). We present here three cases of studies. In experimental study (Miller, 1955) of breakdown in $Ge$ stepwise $p$–$n$ junction was found that measured values of exponents in Miller’s relation were lying in range from 3 to 6.6. The same values of exponents are obtained from expressions (74) and (75) with $\xi = 4$ if we take into account that in $Ge$ with doping levels used in (Miller, 1955) $K_0 \approx 2 \div 3$ (Sze, 1981), (Tsang, 1985), (Miller, 1955), (Grekhov & Serezhkin, 1980), (Stillman & Wolf, 1977), (Shotov, 1958). In experimental study (Bogdanov et al, 1986) of APD based on MIS structure (metal-insulator-semiconductor APD) multiplication of charge carriers occurs in thick $p$–$Si$ substrate. From point of view of avalanche process this structure is similar to asymmetric stepwise $n^+–p$ junction. Therefore, avalanche process in MIS APD can be described by expressions (74)-(76) with $\xi = 4$. Concentration of shallow acceptors in substrate of investigated structure was $10^{15}$ cm$^{-3}$. At this doping avalanche breakdown in $Si$ occurs when electric field near insulator-semiconductor interface reaches value $E_{BD} \approx 3 \times 10^5$ V/cm (Sections 3.1 and 3.2, (Sze, 1981), (Osipov & Kholodnov, 1987), (Sze & Gibbons, 1966)), and therefore $K_0 \approx 10^{-2}$ (Sze, 1981), (Tsang, 1985), (Grekhov & Serezhkin, 1980), (Sze & Gibbons, 1966), (Stillman & Wolf, 1977), (Kuzmin et al, 1975). Measured in (Bogdanov et al, 1986) value $n_n$ at $V_{BD} - V < < V_{BD}$ was found equal to 0.2. From formulas (76) with $K_0 \approx 10^{-2}$ follows that $n_n = 0.186$. In Tables 1 and 2 are presented experimental (Shotov, 1958) and calculated by formulas (48) and (55) values of multiplication factors of electrons $M_n(V)$ and holes $M_p(V)$ in $Ge$ stepwise and linear $p$–$n$ junctions. Obviously, for these $p$–$n$ junctions, experimental and analytical values of multiplication factors differ from each other by less than 20 % in whole voltage $V$ range used in measurements.

<table>
<thead>
<tr>
<th>$V/V_{BD}$</th>
<th>Experiment (Shotov, 1958)</th>
<th>Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.65</td>
<td>1.35</td>
<td>1.30</td>
</tr>
<tr>
<td>0.70</td>
<td>1.50</td>
<td>1.44</td>
</tr>
<tr>
<td>0.75</td>
<td>1.75</td>
<td>1.65</td>
</tr>
<tr>
<td>0.80</td>
<td>2.10</td>
<td>1.98</td>
</tr>
<tr>
<td>0.85</td>
<td>2.65</td>
<td>2.55</td>
</tr>
<tr>
<td>0.90</td>
<td>3.70</td>
<td>3.71</td>
</tr>
<tr>
<td>0.95</td>
<td>7.00</td>
<td>7.30</td>
</tr>
</tbody>
</table>

Table 1. Experimental (Shotov, 1958) and computed [from Equation (48)] hole avalanche multiplication factor $M_p$ in step-wise $p$–$n$ junction in $p$–$Ge$ for different ratios of applied voltage to avalanche breakdown voltage $V / V_{BD}$. It is assumed that $K_0 = 2$ (Shotov, 1958)
Finally, it is interesting to analyze application of expressions (45) and (76) to describe avalanche process in InSb. The fact is that dependence $\alpha(E)$ in InSb was quite well known already in 1967 (Baertsch, 1967), but no one could obtain information about dependence $\beta(E)$ (Dmitriev et al, 1987), (Dmitriev et al, 1983), (Dmitriev et al, 1982), (Gavrjushko et al, 1968). Substituting in (45) dependence $\alpha(E)$ for InSb (Baertsch, 1967), (Dmitriev et al, 1983), (Dmitriev et al, 1982), (Gavrjushko et al, 1968), we find that ratio $K = \beta(E) / \alpha(E)$ is vanishingly small up to electric field $E \cong 4 \times 10^4$ V/cm resulting in extremely high value $n_{\text{eg}}$ when at the same time value $n_{\text{gb}}$ is extremely small. It means that $M_{n}(V)$ becomes much larger than unity, even at voltages $V_b$ noticeably lower avalanche breakdown voltage $V_{BD}$, and value $M_{p}(V)$ remains equal to unity up to values $V_b$ very close to $V_{BD}$. Effect obtained from application of relations (45) and (76) accords very well with experimental data (Baertsch, 1967), (Dmitriev et al, 1983) and explains why multiplication of holes in InSb is extremely hard to observe (Dmitriev et al, 1987), (Baertsch, 1967), (Dmitriev et al, 1983), (Dmitriev et al, 1982), (Gavrjushko et al, 1968).

4. Tunnel currents in avalanche heterophotodiodes

4.1. Calculation of tunnel currents in approximation of quasi-uniform electric field and conditions of its applicability

In act of interband tunneling electron from valence band overcomes potential barrier ABC (Fig. 16a). The length of tunneling $l_T$, i.e. length on which energy of bottom of conduction band $E_c(x)$ changes by value equal to $E_g$ is found by solving integral equation

<table>
<thead>
<tr>
<th>$V/V_{BD}$</th>
<th>$M_p$</th>
<th>$M_n$</th>
<th>$K_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment</td>
<td>Theory</td>
<td>Experiment</td>
<td>Theory</td>
</tr>
<tr>
<td>0.65</td>
<td>1.25</td>
<td>1.19</td>
<td>1.09</td>
</tr>
<tr>
<td>0.70</td>
<td>1.40</td>
<td>1.28</td>
<td>1.14</td>
</tr>
<tr>
<td>0.75</td>
<td>1.60</td>
<td>1.44</td>
<td>1.22</td>
</tr>
<tr>
<td>0.80</td>
<td>1.85</td>
<td>1.70</td>
<td>1.33</td>
</tr>
<tr>
<td>0.85</td>
<td>2.40</td>
<td>2.13</td>
<td>1.70</td>
</tr>
<tr>
<td>0.90</td>
<td>3.50</td>
<td>3.10</td>
<td>2.20</td>
</tr>
<tr>
<td>0.95</td>
<td>6.80</td>
<td>5.89</td>
<td>3.90</td>
</tr>
<tr>
<td>0.975</td>
<td>13.00</td>
<td>11.64</td>
<td>7.00</td>
</tr>
<tr>
<td>0.98</td>
<td>-</td>
<td>14.52</td>
<td>-</td>
</tr>
<tr>
<td>0.985</td>
<td>-</td>
<td>19.33</td>
<td>-</td>
</tr>
<tr>
<td>0.99</td>
<td>30.00</td>
<td>28.90</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2. Experimental (*) (Shotov, 1958) and computed [from Equation (55)] avalanche multiplication factors $M_p$ and $M_n$ for holes and electrons in Ge linear $p-n$ junction for different ratios of applied voltage to avalanche breakdown voltage $V/V_{BD}$ (Shotov, 1958)
\[ E_g = q \times \int_x^{x + l_T(x)} E(x') dx' \]  

(82)

If variation of electric field within length of tunneling \( \Delta E < E \), i.e. specific length of variation of field \( l_E > l_T \), then expanding function \( E(x') \) in Taylor series around point \( x' = x \), we find that in the first order of parameter of smallness \( l_T / l_E \) equation (82) takes the form

\[ l_T = \frac{E_g}{qE(x) \times \left[ 1 - \left( l_T / 2E \right) \times \frac{\partial E}{\partial x} \right]} \]  

(83)

When \( N(x) = \text{const} \) then equation (83) is exact. As can be seen from Fig. 16a, if

\[ |C'C| = \Delta l < l_T, |C'B'| = \Delta E _c \ll E_g, \]  

(84)

then true ABC barrier coincides to high degree of accuracy with triangle ABC' to which corresponds uniform field \( E(x) \) (Fig. 16b).

It follows from (83) and Poisson equation that inequalities (84) are satisfied if

\[ \bar{\delta}(x) = \frac{N(x) \times E_g}{2EE_0 \times E^2(x)} \ll 1, \]  

(85)

at that

\[ l_T(E_g, E) = \frac{E_g}{q \times E(x)} \]  

(86)

As shown below, due to large values of field \( E \) at avalanche breakdown of \( p-n \) structures, inequality (85) is valid for almost all materials up to concentration \( N = 10^{17} \text{ cm}^{-3} \) and even high.

Under these conditions specific rates of charge carriers’ tunnel generation \( g_Ti(x) \) in layers I and II of structure can be described by expression

\[ g_Ti(x) = \frac{1}{q} \times \frac{\partial j_Ti}{\partial x} = A_{gi} \times E^2(x) \times \exp \left[ -\frac{a}{E(x)} \right], \]  

(87)
obtained in (Kane, 1960) (see also (Burstein & Lundqvist, 1969)) for $E(x) = \text{const}$, in which

$$v = \frac{p}{m_i} = \frac{2m_i \times m_i}{(m_i + m_c) - \text{crossed Plank constant, gaps and specific effective masses of light charge carriers in proper layers. Approximation of quasi-uniform field (87)}}$$

Here $\hbar$, $E_g$, and $m_i = \frac{2m_i \times m_i}{(m_i + m_c) - \text{crossed Plank constant, gaps and specific effective masses of light charge carriers in proper layers. Approximation of quasi-uniform field (87)}}$
and expressions (6)-(9) result in convenient formula for analysis of primary interband tunnel current density

\[ J_T = \sum_i J_{Ti} = \sqrt{2} \frac{q^3}{(2\pi)^3} \rho^2 \times \sum_i \left( \frac{n_i}{E_{g i}} \times L_{Ti} \times E_i^2 \times \exp \left( -\frac{a_i}{E_i} \right) \right), \]

(89)

where characteristic dimensions of areas of charge carriers’ tunnel generation in layers I and II

\[ L_{Ti}(E_i, W_i) = \min \left\{ W_{Ti} \equiv \varepsilon_0 \varepsilon_i \times E_i^2 \times q \times a_i \times N_i, W_i \right\}. \]

(90)

Equation (89) is valid under conditions

\[ \frac{N_i \times E_{gi}^2}{2\varepsilon_0 \times E_i^2} < \frac{E_i}{a_i} << 1, \]

(91)

\[ I_{Ti} = I_T(E_{gi}, E_i) = \frac{E_{gi}^2}{q \times E_i} << I_j. \]

(92)

These conditions mean the following. If inequalities (91) for \( g_{Ti}(E) \) are satisfied then expression (87) is valid, at least in the neighborhood of field value \( E = E_i \). When right side of inequalities (91) is satisfied then tunnel generation drops sharply with decreasing \( E_i \), and therefore \( I_{Ti} \) at \( W_{Ti} < W_i \) is mainly determined by tunneling in areas \( 0 \leq x < W_{T1} \) and \( W_1 \leq x \leq W_1 + W_{T2} \).

Fulfillment of conditions (92) is necessary at punch-through of proper layers of structure for neglecting tunneling through its hetero-interfaces which is not accounted for by formula (89). We show further, that at avalanche breakdown, inequalities (91) and (92) are valid for almost all real values of material parameters, concentrations \( N_i \) and layers’ thicknesses \( W_i \) of heterostructure. Avalanche breakdown occurs when one of fields \( E_i \) becomes close to breakdown field \( E_{iBD} \) of proper layer of structure ((Sze, 1981), (Tsang, 1985), (Grekhov & Serezhkin, 1980), Sections 3.1-3.3).

Breakdown fields \( E_{iBD} \) can be obtained by formula (14) (Osipov & Kholodnov, 1987), (Osipov & Kholodnov, 1989), i.e.,

\[ E_{iBD}(N_i, W_i) = E_{iBD}(0, W_i) \times \left[ 1 + \frac{N_i}{N_i(W_i)} \right]^{-1/s}, \]

(93)
where

\[ E_{BD}(0,W) = A_i \times \left( \frac{A_i \times \varepsilon_i \varepsilon_0}{s_0 W_i} \right)^{1/3}, \quad \tilde{N}_i(\tilde{W}) = \left( \frac{A_i \times \varepsilon_i \varepsilon_0}{s_0 W_i} \right)^{1/3} \]  

(94)

(s and \( A_i \) – some constants).

For many semiconductors including \( In_x Ga_{1-x} As \) alloy which is one of the main materials for avalanche heterophotodiodes fabrication (Tsang, 1981), (Stillman, 1981), (Filachev et al, 2010), (Kim et al, 1981), (Forrest et al, 1983), (Tarof et al, 1990), (Ito et al, 1981), (Clark et al, 2007), (Hayat & Ramirez, 2012), (Filachev et al, 2011), (Stillman et al, 1983), (Ando et al, 1980), (Trommer, 1984), (Woul, 1980),

\[ s = 8, \quad A_i = \sqrt{\frac{12 \times \varepsilon_i \varepsilon_0}{\varepsilon_i \varepsilon_0}} \times \left( \frac{E_{gi}}{11q} \right)^{3/4} \times 10^{10}. \]  

(95)

From expressions (93) and (94) when relations (95) are satisfied we find the following.

1. When

\[ N_i \leq N_i^{(1)} = \frac{8.9 \times 10^{19}}{X_{mi}^4 \times X_{ci}^4 \times X_{gi}^6}, \quad \text{cm}^{-3}, \quad W_i \geq W_i^{(1)} = X_{mi}^{3.5} \times X_{ci}^2 \times X_{gi}^6 \times 1.4 \times 10^{-4}, \mu m, \]  

(96)

then ratio \( E_i \) to \( a_i \) is less than 0.1, where \( X_{mi} = 0.06 / m_i^* \), \( X_{ci} = 12.4 / \varepsilon_i \), \( X_{gi} = 1.35 / E_{gi} \) (for InP which is often used for growing of wide-gap layers of heterostructure (Tsang, 1981), (Stillman, 1981), (Filachev et al, 2010), (Kim et al, 1981), (Forrest et al, 1983), (Tarof et al, 1990), (Ito et al, 1981), (Clark et al, 2007), (Hayat & Ramirez, 2012), (Filachev et al, 2011), (Stillman et al, 1983), (Ando et al, 1980), (Trommer, 1984), (Woul, 1980)), \( X_{mi} = X_{ci} = X_{gi} = 1, m_0 = m_i^* / m_0 \) (\( m_0 \)– free-electron mass)

2. When

\[ N_i \leq N_i^{(2)} = X_{mi}^{0.2} \times X_{ci}^{1.6} \times X_{gi}^{0.4} \times 3.3 \times 10^{17}, \text{cm}^{-3}, \quad W_i \geq W_i^{(2)} = \frac{X_{mi}^{0.4} \times 1.8 \times 10^{-2}}{X_{mi}^{0.2} \times X_{ci}^{1.9}}, \mu m \]  

(97)

then under avalanche breakdown of proper layer of structure ratio \( \delta_i \) to \( E_{BD} / a_i \) is not exceed unity, moreover, even when \( N_i = N_i^{(2)} \)

\[ \delta_i < X_{mi}^{0.6} \times X_{ci}^{1.2} \times X_{gi}^{0.8} \times 10^{-1}. \]  

(98)
When

\[ W_i > > \frac{1.8 \times 10^{-2}}{\sqrt{X_{1i} \sqrt{X_{gi}}}} \, \mu m, \]  

(99)

then length of tunneling \( l_{Ti} \) at \( E_i = E_{iBD} \) is much shorter than thickness \( W_i \) of this layer.

In expressions (96)-(99) \( E_{gi} \) is measured in eV. Analysis shows that under avalanche breakdown of heterostructure inequities (91) and (92) are satisfied for real values of \( N_i \) and \( W_i \) and \( E_i < E_{iBD} \) i.e. in layer which does not control avalanche breakdown also. As can be seen from Fig. 17, when punch-through of layer \( n_{wg} \) stops then, obviously, conditions (91) and (92) become no longer valid. Note that calculations of tunnel currents in approximation of quasi-uniform field lead to some overestimation of actually available. In fact, due to high doping of \( p^{+} \) layer, tunnel current in it can be ignored; this is situation similar to MIS structures (Anderson, 1977). In \( n \) type layers electric field decreases with increasing distance from metallurgical boundary of \( p^{+}-n \) junction (Fig. 1b), and because gradient of potential is expressed as \( \frac{d\phi}{dx} = -E \) then slope of zones \( E_c(x) \) and \( E_v(x) \) decreases with increasing \( x \). It is shown from Figure 16a that use of quasi-uniform field approximation means underestimating of thickness of actual barrier ABC. As expected, numerical calculations in WKB approximation (Anderson, 1977) give a somewhat smaller value of tunnel currents than formula (89). Since tunnel currents are strongly dependent on parameters of material, which in real samples, usually, more or less different from those used in calculations (moreover, exact dopant’s distribution profile \( N_i(x) \) and hence shape of barrier ABC are usually unknown), then slight overestimation of tunnel currents values provides some technological margin that is needed for development of devices with required specifications.

4.2. Features of interband tunnel currents in \( p^{+}-n \) heterostructures under avalanche breakdown

Analysis of expression (89) under avalanche breakdown of \( p^{+}-n \) heterostructure, i.e., when either \( E_1 = E_{1BD} \) or \( E_2 = E_{2BD} \), shows that in contrast to homogeneous \( p-n \) junction (Stillman, 1981), (Ando et al, 1980) density of initial tunnel current \( J_{Ti} \), as a rule, is not a monotonic function \( N_2 \). An increase in \( N_2 \) cause, for some values of \( N_1 \) and \( W_i \), the rise of tunnel current and vice versa – decrease of tunnel current when \( N_1 \) and \( W_i \) have different values. Depending on gap \( E_{gi} \) of heterostructure’s layers and their thicknesses \( W_i \) the following situations are possible.

4.2.1. Independent doping levels of wide-gap and narrow-gap \( n \) type layers

I.

\[ \frac{W_1}{W_2} = W_{1/2}^2 = \left( \frac{N_1}{N_2 + N_2(W_2)} \right)^{s-1/2} \]  

(100)
In this case, at any concentration \( N_1 \), field \( E_1 = E_{1BD}(N_1, W_1) \), and \( E_2 < E_{2BD} \), i.e., avalanche breakdown is controlled by \( n_{wg} \) layer.

As follows from (6)-(9), (89) and (93), if

\[
\exp\left[-\frac{a_1}{E_{1BD}(0, W_1)} \times \left(1 - \frac{\varepsilon_2 \times a_2}{\varepsilon_1 \times a_1}\right)\right] < 1,
\]

(101)

which is fulfilled with large margin at \( a_2 \varepsilon_2 < a_1 \varepsilon_1 \) due to large ratio of \( a_1 \) to \( E_{1BD}(0, W_1) \) (1-2 orders of magnitude) while

Figure 17. Dependence of generalized parameters of smallness \( \delta_2^* \) and \( l_T^* \) in quasi-uniform field approximation on concentration \( N_1 \), at \( M_{ph} = 100 \), in case, when charge carriers multiplication occurs in \( n_{wg} \):InP layer. Solid lines – \( \delta_2^* \), dashed – \( l_T^* \). Values \( W_1 \) : 1 - 0.5, 2 - 2, 3 - 8. \( N_{1pt} \) - maximal concentration \( N_1 \), at which punch-through of \( n_{wg} \) layer is possible; \( \delta_2 = (N_2 / 10^{16}) \times (\varepsilon_1 / \varepsilon_2) \times E_{g2} \times \delta_2^*; l_T^* = (\varepsilon_1 / \varepsilon_2) \times E_{g2} \times l_T^* \) eV, concentration – cm\(^{-3}\).
\[
N_1 < \tilde{N}_1^{(T)} \equiv s \times \left( \frac{2}{s-1} \times \frac{e_1}{e_2 \times a_2} \times E_{1BD}(0, W_1) \right)^{1/2} \times \tilde{N}_1(W_1) \times W_1^{(s+0.5)/(s-1)}
\]

then tunnel current is almost independent on \( N_1 \).

If \( s \) sufficiently large (Sze, 1981), (Osipov & Kholodnov, 1987), (Sze & Gibbons, 1966), Sections 3.1-3.3), then with further increase of \( N_1 \) tunnel current is monotonically falling. However, in most real cases, for example, when relations (95) is valid, tunnel current at \( N_1 > \tilde{N}_1^{(T)} \) first decreases and then increases.

One can see that at minimum of tunnel current, as a rule, the following inequality is valid

\[
\xi \equiv \frac{E_{1BD}(0, W_1)}{a_1} \times \frac{\kappa^{(s-2)/(s-1)}}{s^{1/(s-1)}} \times \frac{y}{f^2(y)} \times \frac{s}{r \times N_1}
\]

\( f(y) = (y + r^{-1})^{1/s}, r = (\kappa \times s)^{1/(s-1)}, \kappa = 1 - \frac{a_2 \times e_2}{a_1 \times e_1}, y = \frac{N_1}{r \times N_1} \)

When (103) is fulfilled then \( W_{T1} < W_1 \).

Therefore, as it follows from (6)-(9), (89), (90) and (93), concentration \( N_1 = N_{1\text{min}} \) at which \( J_T \) reaches minimum is defined by equation

\[
\frac{y}{f(y)} + \frac{\xi}{r^{1/(2/s)}} \times \left[ s \times f(y) - r^{(1/s)} \times y \right] \times \ln[\Lambda(y; \xi)] = 1,
\]

\( \Lambda(y; \xi) = B \times \frac{f^{3/(s \times (y))}}{y \times [f(y) - \kappa \times y] \times [\kappa \times s - f^{1/(s \times (y))}]} \times \left[ 1 - \frac{\kappa^{1/s \times (y)} \times f(y) \times \left( s - 4 + \frac{s}{r \times y} \right)}{1 + \frac{4r}{(1 - \kappa) \times \kappa \times s} \times f(y) - \kappa \times y} \right] \times \frac{s}{r \times N_1} \)

\( B = \left( \frac{m_1}{m_2} \right)^{3/2} \times \left( \frac{E_{x1}}{E_{x2}} \right)^{3/2} \times \frac{N_2}{N_1} \times \frac{1 - \kappa^2}{r} \times \frac{1}{r} \)

Expression (105) is valid when inequality \( W_{T2} < W_2 \) is fulfilled. This inequality and inequality (103) also are fulfilled at minimum of tunnel current in the most practically interesting cases. Below is explained difference between situations \( W_{T2} > W_2 \) and \( W_{T2} < W_2 \) at
\( N_1 = N^{(T)}_{\text{min}} \). Equation (104) can be solved by successive approximations using parameters of smallness \( \xi \) and \( 1/s \).

As a result we find

\[
N^{(T)}_{\text{min}} = \left[ \frac{\varepsilon_{\text{eff}} \times A_1}{q \times W_1} \left( 1 - \frac{\varepsilon_2 \times d_2}{\varepsilon_1 \times d_1} \right) \right]^{\frac{1}{2}} \times y_0 \times \left( 1 - \frac{1 - \kappa}{\kappa} \right) \times e^{\frac{1}{2}} \times \ln \left[ \Lambda(y_0;0) \right] \times \frac{y_0 \times (\kappa \times s \times y_0 + 1)}{(s-1) \times \kappa \times y_0 + 1} + 0(1), \quad (106)
\]

where

\[
y_0 = 1 + \frac{1}{\kappa \times s^2} + 0 \left( \frac{1}{s^2} \right). \quad (107)
\]

It is shown from (105) and (106) that \( N^{(T)}_{\text{min}} \) is decreased with growth \( W_1 \) and, also, although weakly, with increase \( N_2 \).

When \( N_1 = N^{(T)}_{\text{min}} \), then density of tunnel current

\[
I_T(N_1) = I_{T\text{min}} = C_0 \times \frac{\varepsilon_{\text{eff}} \times q^2}{2\pi \times h \times E_{\text{g}}} \times \frac{E_{\text{g}}^4(0,W_1)}{N_1(W_1)} \times \Lambda^{-n_1}(y_0;0) \times \exp \left[ -\frac{C_1 \times a_1}{E_{\text{g}}(0,W_1)} \right] \times [1 + 0(1)], \quad (108)
\]

Where

\[
C_0 = y_0^3 \times \frac{y_0 \times \kappa \times (s-1) + 1}{(s-1) \times \kappa \times y_0 + 1} \times (\kappa \times s)^{1/(s-1)}, \quad C_1 = \left[ y_0 \times (\kappa \times s)^{1/(s-1)} \right]^{-1},
\]

\[
n_1 = \frac{y_0 \times (1-\kappa)}{(s-1) \times \kappa \times y_0 + 1}.
\]

From (94), (105) and (108) follow that \( I_{T\text{min}} \) decreases sharply with increasing \( W_1 \). Value \( I_{T\text{min}} \) decreases also, although weakly, with increasing \( N_2 \). Ratio

\[
\frac{I_{T\text{min}}}{I_T(N_1)_{N_1 < N^{(T)}_{\text{min}}}} \propto \left[ \frac{N_2}{N_1(W_1)} \right]^{n_2} \times \exp \left[ -(C_1 + \kappa - 1) \times \frac{a_1}{E_{\text{g}}(0,W_1)} \right], \quad (109)
\]

Where

\[
n_2 = \frac{y_0 \times (\kappa \times s - 1) + 1}{(s-1) \times \kappa \times y_0 + 1}.
\]
drops sharply, same as $I_{r\min}$, with increase $W_2$, but it increases with increasing $N_2$. Value of this ratio is usually several orders of magnitude less than unity. For example, for combination of layers $n_{wg}:InP/n_{ng}:n_{0.53Ga_{0.47}As}$, differential of currents, as can be shown, does not exceed values $(N_2/10^{18})^{0.9} \times 2 \times 10^{-4}$, where $N_2$ is measured in cm$^{-3}$.

When concentrations

$$N_2 < \frac{\varepsilon_2 \times a_2}{\varepsilon_1 \times a_1 - \varepsilon_2 \times a_2} \times \frac{W_1}{W_2} \times N_{1T_{min}} \quad (110)$$

then in minimum of $I_T(N_1)$ takes place punch-through of narrow-gap layer, i.e. non-equilibrium SCR reaches $n_{ng}$ layer. When $N_1 > N_{1T_{min}}$, tunnel current increases with increasing $N_1$, and at the same time, non-equilibrium SCR will penetrate into narrow-gap layer until concentration $N_1$ reaches value

$$N_1 = N_{1T_{min}} \left( A_1 \times \frac{\varepsilon_1 E_0}{q W_1} \right)^{\nu/(\nu - 1)} \times (1 + 0(1)) > N_{1T_{min}} \quad (111)$$

Nature of above dependence $I_T$ on $N_1$ is competition between tunnel currents in wide-gap and narrow-gap layers of heterostructure (Fig. 1a). When $N_1 \leq N_{1T}$ then field $E = E_i(W_1)$ in $n_{wg}$ layer at its heterojunction (Fig. 1b) coincide with very high accuracy with $E_{1BD}$. Due to relatively large field $E_2 = (\varepsilon_1/\varepsilon_2) \times E_{1BD}$, current density $I_T$ is determined by tunneling of charge carriers in narrow-gap layer, i.e. $I_T = I_{T2}$ (Fig. 1a). With increasing $N_1$, field $E_2$ and therefore current $I_{T2}$ decrease due to fall $E_i(W_1)$ (Fig. 18). Decrease $E_i(W_1)$ with increase $N_1$ is caused by requirement (1) of constancy of photocurrent gain $M_{ph} = M_r$. Indeed, increase $N_1$ for given $M_{ph}$ should lead to growth $E_1$. Otherwise, due to growth $\nabla E(x)$ with increasing $N_1$, field would be reduced everywhere in SCR, which in turn would lead to a decrease $M_{ph}$. However, increase $E_1$ should not be too large, and it should be such that $E(x)$ at $x > W_1$ is decreased. In other words, $E(x)$ anywhere in SCR would increase, that, evidently, would increase $M_{ph}$. It can be seen directly from (1) and (2). Note that for sufficiently large values of multiplication factors $M_{ph}$, field $E_1$ is practically independent on $M_{ph}$ and very close to breakdown field $E_{1BD}(N_1, W_1)$ when value of integral $m$ (2) is equal to unity. This allows to use value $E_1 = E_{1BD}(N_1, W_1)$ (93) instead of true value $E_1(N_1, W_1, M_{ph})$. When $N_1 > N_{1T}$, then variation of field $E(x)$ at distance $W_1$ in $n_{wg}$ layer is still very insignificant, but it is enough to affect value $I_{T2}$. Due to decrease $E_2$ with growth $N_1$ (especially when $N_1 > N_{1T}$), current is more and more determined by tunneling of charge carriers in $n_{wg}$ layer, therefore when $N_1 > N_{1T_{min}}$, current density $I_T = I_{T1}$ in-
creases with increase $N_1$ because $E_{1BD}$ grows with increase $N_1$. Initial plateau (Fig. 18a) on the graph $J_T(N_1)$ is caused by extremely weak dependences $E_{1BD}$ on $N_1$ (93) and $E$ on $x$ in $n_{wg}$ layer when $N_1<\bar{N}_1(T)$. Reducing of value $J_{T \text{min}}$ (108) with growth $N_2$ is due to increasing length of tunneling in narrow-gap $n_{ng}$ layer (Fig. 1). Indeed, in this layer $\nabla E = -N_2<0$, and $E_2$ under these conditions does not depend on $N_2$. It means, that $E(x)$ everywhere in $n_{ng}$ layer, except of point $x=W$, falls with increasing $N_2$ (1b). Since $\frac{dE_c}{dx} = \frac{dE_v}{dx} = \frac{d\phi}{dx} = -E < 0$, then slopes of $E_c(x)$ and $E_v(x)$ everywhere in $n_{ng}$ layer, except of point $x=W$, decrease also with increasing $N_2$, that leads to increase length of tunneling. Reducing of $J_T$ is more significant with growth $N_2$ when $N_1<N_{1\text{min}}$ (Fig.18b), because current density $J_{T2}$ increases with decrease $N_1$ while $J_{T1}$ decreases. When $N_1=N_{1\text{min}}$ then current density $J_{T1} \leq J_{T2}$ and if $N_1=N_{1\text{min}}$ it exceeds $J_{T\text{min}}$. Therefore, ratio of $J_{T\text{min}}$ to $J_T\mid_{N_1<N_{1\text{min}}}$ (109) increases with increasing $N_2$. Because at $N_1=N_{1\text{min}}$ value $J_{T1}>J_{T2}$, then, naturally, concentration $N_{1\text{min}}$ (106) slightly decreases with increasing $N_2$ (Fig. 18b). For small values $N_2$, when $W_{T2}>W_2$, $E(x)$ in $n_{ng}$ layer coincides with $E_2$ with high accuracy. Therefore, length of tunneling in this layer, and hence $J_T$ also, do not depend on $N_2$. Reducing of values $N_{1\text{min}}$ (106) and $J_{T\text{min}}$ (108) with increasing $W_1$ (Figure 18a) is due to the fact that the more is $W_1$ then the less is $E_{1BD}$ and the greater is fall of field $E(x)$ in depth of $n_{ng}$ layer.

II. Condition (100) is not satisfied. For example, for combination of layers $n_{wg}$:InP / $n_{ng}$:In$_{0.53}$Ga$_{0.47}$As such situation takes place when

$$\frac{W_1}{W_2} \left[1 + \frac{N_2}{2.2 \times 10^{15}} \times W_2^{8/7}\right]^{7/8} < 21.5,$$

(112)

where $N_2$ and $W_2$ are measured in cm$^{-3}$ and μm, respectively. Under this condition, when $N_1<\bar{N}_1$, where $\bar{N}_1$ satisfies equation

$$\frac{E_s}{e_1} \times A_2 \times \left[N_2 + \bar{N}_2(W_2)\right]^{3/8} + \frac{q \times \bar{N}_1 \times W_1}{\epsilon_1 e_0} = A_1 \times \left[N_1 + \bar{N}_1(W_1)\right]^{1/8},$$

(113)

avalanche breakdown is controlled by $n_{ng}$ layer, i.e. $E_2 = E_{2BD}(N_2, W_2)$, and $E_1 < E_{1BD}$ and it increases linearly with $N_1$. Therefore, strictly speaking, when $N_1<\bar{N}_1$ then tunnel current increases with increasing $N_1$. At the same time, $J_{T2}$ does not depend on $N_1$ under following conditions.
Figure 18. Dependence of tunnel current density on concentration $N_1$ in case of independent doping levels of $n_{wg}$: InP and $n_{ng}$: $In_{0.53}Ga_{0.47}As$ layers at $W_2 = 2 \mu m$. 

a - $N_2 = 10^{14} cm^{-3}; W_1, \mu m: 1 - 0.1, 2 - 0.2, 3 - 0.5, 4 - 1$. 

b - neighborhood of value $N_1 = N_{1min}(T)$. 

$W_2 = 2 \mu m; N_2, cm^{-3}: 1 - 10^{14}, 2 - 10^{15}, 3 - 10^{16}, 4 - 10^{17}$
1. If
\[
\left( \frac{W_{1/2}}{W_{1/2}} \right)^{1/(s-1)} > 1 - \frac{s-1}{2s^2},
\]
then at \( N_1 < \bar{N}_{1/2} \) with margin of several orders of magnitude, and therefore with very high accuracy \( J_T(N_1) = \text{const} \). If \( N_1 > \bar{N}_1 \) then due to decrease \( E_2 \) and hence \( J_{T2} \) also, density of tunnel current \( J_T(N_1) \) begins drop sharply and, reaching minimum value (108) at concentration (106), then starts to grow again due to growth \( J_{T1}(N_1) \).

2. If
\[
\left( \frac{W_{1/2}}{W_{1/2}} \right)^{1/(s-1)} << 1 - \frac{s-1}{2s^2},
\]
then after initial plateau \( J_T(N_1) \) grows monotonically. It is due to monotonic increase in component of tunnel current density \( J_T(N_1) \), which at \( N_1 \geq \bar{N}_1 \) is considerably superior to \( J_{T2} \).

3. If
\[
\left( \frac{W_{1/2}}{W_{1/2}} \right)^{1/(s-1)} \approx 1 - \frac{s-1}{2s^2},
\]
then for small enough thicknesses \( W_1 \) of layer \( n_{wg} \), dependence \( J_T(N_1) \) has distinct maximum at \( N_1 = \bar{N}_1 \), however, at least in this case minimum is not deep. This is due to the fact that components of tunnel current density \( J_{T1} \) and \( J_{T2} \) are equal to each other in order of magnitude at small enough \( W_1 \). Characteristics of tunnel currents in heterostructure with independent doping of \( n_{wg} \) and \( n_{ng} \) layers are illustrated in Fig. 18. Note that if in case I increase \( N_2 \) leads to decrease \( J_T \) at all values \( N_1 \) then in case II, increase \( N_2 \), when \( N_1 \) is small enough, leads to increase of tunnel current, but at sufficiently large \( N_1 \) tunnel current decreases, particularly, in the vicinity of concentration \( N_1 = N_{1min}^{(T)} \).

4.2.2. Equal doping levels of wide-gap and narrow-gap \( n \) type layers

Under this condition density of tunnel current is given by expression (89), where \( N_1 = N_2 = N \). The
\[
\frac{W_1}{W_2} \geq \left( \frac{E_1 \times A_1}{E_2 \times A_2} \right)^e
\]
At this relation of parameters avalanche breakdown is controlled by \( n_{wg} \) layer, i.e. \( E_1 = E_{1BD}(N_1, W_1) \), and \( E_2 = E_{2BD}(N_2, W_2) \) regardless of doping. Dependence \( J_T \) on \( N \) has identi-
cal character with $J_T(N_1)\left|_{N_2 = \text{const}}\right.$ in the case of 4.2.1. $I$, and is caused by the same physical grounds. The only difference is that when $N < N_2$, then curves $J_T(N)$ lie higher on plotting area, and when $N > N_2$—lower, than curves $J_T(N_1)\left|_{N_2 = \text{const}}\right.$ in the case of 4.2.1. $I$. This occurs because at given value $E_2$ length of tunneling in narrow-gap layer is the greater the higher is level of doping of this layer.

ii. Condition (117) is not satisfied. Then, till $N < \bar{N}$, (where $\bar{N}$ is determined by equation (113), where $\bar{N}_1=\bar{N}_2=N$) avalanche breakdown is controlled by $n_{wg}$ layer, i.e. $E_2=E_{2\text{BD}}(N, W_2)$, and $E_1<E_{1\text{BD}}(N, W_1)$ and increases linearly with $N$. Dependence $J_T(N)$ has, in contrast to situation 4.2.1, not only deep minimum, but high maximum also (Fig. 19a). This is due to the fact that when $N < \bar{N}$ then $E_1$ grows and $E_2$ grows also reaching at $N = \bar{N}$ maximal value (Fig. 19b). As a result, when $N < \bar{N}$ then $J_{T1}$ grows with increase $N$ and $J_{T2}$ grows also. Note that when doping of $n_{wg}$ and $n_{ng}$ layers are equal then concentration $N = N_{\text{min}}^{(T)}$ at which tunnel current density $J_T$ has minimal value, is determined by formula (106) with accuracy up to small corrections of order $\xi=E_{1\text{BD}}(0, W_1)/a_1 < < 1$, as in the case of independent doping of $n_{wg}$ and $n_{ng}$ layers. Formula for $J_T^{\text{min}}$ may be obtained from expression (108), if we replace $N_2$ by $N_{\text{min}}^{(T)}$ in it.

5. Basic performance of avalanche heterophotodiode

5.1. Responsivity

In punch-through conditions of absorber $n_{ng}$, current responsivity $S_J(\lambda)$ of heterostructure under study can be described by relation (4). In calculating quantum efficiency $\eta$ of heterostructure, we take into account that optical radiation is not absorbed in its wide-gap layers. Let’s assume that light beam falls perpendicularly to front surface of heterostructure (Fig. 1), and absorption coefficient in narrow-gap layer $\gamma(\lambda)$ does not depend on electric field. Quantum efficiency is ratio of number of electron-hole pairs generated in sample by absorbed photons per unit time to incident flux of photons.

Therefore, (Fig. 20a)

$$\eta = \frac{(1-R_1) \times (1-R_2)}{1-R_1 \times R_2} \times \eta_t,$$

where reflection coefficient of light from illuminated surface $R_1 = \left(\sqrt{\varepsilon_{ex}} - \sqrt{\varepsilon_1}\right)^2 / \left(\sqrt{\varepsilon_{ex}} + \sqrt{\varepsilon_1}\right)^2$ and from interfaces of heterostructure $R_2 = \left(\sqrt{\varepsilon_2} - \sqrt{\varepsilon_1}\right)^2 / \left(\sqrt{\varepsilon_2} + \sqrt{\varepsilon_1}\right)^2$; $\varepsilon_{ex}$—relative dielectric
Figure 19. Dependences of tunnel current density (a) and fields $E$ (b) on dopant concentration $N$ in case of equal doping levels of $n_{\text{wg}} : \text{InP}$ and $n_{\text{ng}} : n_{0.53}\text{Ga}_{0.47}\text{As}$ layers, at $W_2 = 2$ μm. $W_1$, μm: 1 – 10, 2 – 1, 3 – 0.1, Curves 1', 2', 3' – $E_2(N)$; curve 4 – $E_1(N)$, weakly dependent on $W_1$. 

Photodiodes - From Fundamentals to Applications
constant of environment; and quantum efficiency $\eta_1$ with respect to light ray which has penetrated into narrow-gap layer is written

$$\eta_1 = 1 - \zeta + \eta_2 \times \zeta;$$

(119)

quantum efficiency $\eta_2$ with respect to light ray which has reached to second interface of heterostructure,

$$\eta_2 = R_2 (1 - \zeta) + \frac{R_2 \zeta (1 - \zeta) + \eta_2 \zeta R_3^2 + \frac{R_1 R_2 (1 - R_3)^2 \zeta}{1 - R_1 R_2}}{1 - R_1 R_2} (1 - \zeta + \eta_2 \zeta) + \frac{(1 - R_3)^2 R_3}{1 - R_3 R_2} \times$$

$$\times \left[ (1 - \zeta)(1 + R_2 \zeta) + \eta_2 R_2 \zeta^2 + \frac{R_1 (1 - R_3)^2 \zeta}{1 - R_1 R_2} (1 - \zeta + \eta_2 \zeta) \right];$$

(120)

$\zeta = \exp(-\gamma W_2)$, $R_3$ – reflection coefficient of light from not illuminated (backside) surface.

From expressions (118)-(120) follow, that

$$\eta(\gamma W_2) = \eta(\infty) \times \left[ 1 - \exp(-\gamma W_2) \right] \times \frac{1 + R_{23} \exp(-\gamma W_2)}{1 - R_{12} R_{23} \exp(-2\gamma W_2)}$$

(121)

where

$$\eta(\infty) = \frac{(1 - R_1)(1 - R_2)}{1 - R_1 R_2},$$

(122)

$$R_{ij} = \frac{R_i (1 - R_j) + R_j (1 - R_i)}{1 - R_i R_j}, i, j = 1, 2, 3.$$  

(123)

Particularly,

$$\eta(\gamma W_2) = \begin{cases} 
\frac{\eta(\infty) \exp(-\gamma W_2)}{1 - R_{12} \exp(-\gamma W_2)}, & \text{at } R_3 = R_1, \\
\frac{\eta(\infty) \exp(-2\gamma W_2)}{1 - R_{12} \exp(-2\gamma W_2)}, & \text{at } R_3 = 1. 
\end{cases}$$

(124)

Dependence $\eta$ on $W_2$ for heterostructure $\text{InP} / \text{In}_{0.53}\text{Ga}_{0.47}\text{As} / \text{InP}$ is shown in Fig. 20b. It should be noted that since in operation, electric field is high even in absorption layer, then,
due to Franz-Keldysh effect, quantum efficiency is slightly higher than given in Fig. 20b. This is especially true when absorbing layer $W_2$ is very thin.

**Figure 20.** Layout view of multiple internal reflections and absorptions of light beam in heterostructure (a) and dependence of quantum efficiency $\eta$ of structure $\text{InP} / \text{In}_{0.53}\text{Ga}_{0.47}\text{As} / \text{InP}$ on absorption layer thickness $W_2$, $\mu$m (b): $1 - R_3 = R_1$, $2 - R_3 = 1$. It is assumed that relative dielectric permittivity of environment $\epsilon_{\text{ex}} = 1$
5.2. Noise

It was noted above that in order to achieve the best performance of SAM-APD special doping profile is formed in heterostructure which facilitates penetration of photogenerated charge carriers with higher impact ionization coefficient into multiplication layer. In this case, at given voltage bias on heterostructure, current responsivity 
\[ S_I(\lambda) \]
is maximal, and effective noise factor 
\[ F_{ef}(M_{ph}) \]
is minimal (Tsang, 1985), (Filachev et al, 2011), (Artsis & Khododnov, 1984), (McIntyre 1966), and hence, as it is evident from expression (5), noise spectral density 
\[ S_N \]
is also minimal. If \( \alpha = \beta \), then (Tsang, 1985), (Filachev et al, 2011), (Artsis & Khododnov, 1984), (McIntyre 1966) 
\[ F_{ef}(M_{ph}) = M_{ph} \]
and therefore
\[ S_N = 2q \times A \times M_{ph}^2. \]

In InP ratio \( K(E) = \beta / \alpha \) in interval of fields of interest \( E = (3.3 \div 7.7) \times 10^5 \) V/cm varies from 2.3 to 1.4 (Tsang, 1985), (Filachev et al, 2011), (Cook et al, 1982). Therefore, noise spectral density of heterostructure with InP multiplication layer and optimal doping is slightly less than value given by formula (125). When \( N_1 > \bar{N}_1 \) (where \( \bar{N}_1 \) satisfies equation (113) (see Fig. 21), in which \( \bar{N}_i(W_i) \) is defined by formula (94) for \( i = 1, 2 \) then avalanche multiplication of charge carriers in narrow-gap layer does not occur. Under these conditions, field value at metallurgical boundary of \( p^- - n \) junction \( (x = 0, \text{Fig. } 1) \) equals to \( E_1 = E_{1b}(N_{1b}, W_1) \) (see (93) and (94)). For many semiconductors (see Sections 3.1-3.2) including \( In_{0.53}Ga_{0.47}As / InP \), values \( s \) and \( A_i \) are defined by relations (95). In the case of heterostructure \( InP / In_{0.53}Ga_{0.47}As / InP \), in first approximation in parameters of smallness
\[ \delta_1 = \frac{E_{1b}(0, W_1)}{a_1} = \frac{2.786 \times 10^{-2}}{W_1^{1/2}}, \quad \delta_2 = \frac{1}{s^2} = \frac{1}{64}, \]
we find that value of concentration \( N_1 = N_{1min}^{(T)} \) at which function \( J_{T}(N_1) \) reaches its minimum
\[ J_{Tmin}(W_1, N_2) = 2.19 \times 10^8 \times \frac{W_1^{0.49}}{N_2^{0.07}} \times \exp(-27.88 \times W_1^{1/7}), \quad A/cm^2, \]
is given by
\[ N_{1min}^{(T)}(W_1, N_2) = \frac{2.33 \times 10^{16}}{W_1^{1/2}} \times \left[ 1 - \frac{2.52 \times 10^{-2}}{W_1^{1/2}} \times \left( \frac{\ln \left( \frac{N_2 \times W_1^{1/7}}{3.69 \times 10^{35}} - 1.41 \right)}{6.49 \times 10^{35}} \right) \right], \quad \text{cm}^{-3}. \]

Formulas (127) and (128) are valid when \( W_{T2} \leq W_2 \), i.e., as follows from Section 4.2.1, when
\[ N_2 \times W_2 \geq Q(W_1) = \frac{5}{W_1^{1/2}} \times 10^{14}, \]
where concentration and thicknesses, as in (127) and (128), are measured in cm\(^{-3}\) and μm, respectively.

If inequality (129) is not satisfied, then values \(N_{1\min}^{(T)}\) and \(J_{T\min}\) will be again determined by (127) and (128), in which \(N_2\) is replaced by \(Q(W_1)/W_2\). It is shown from (127) and (128) that \(N_{1\min}^{(T)}\) and \(J_{T\min}\) are decreasing, moreover \(J_{T\min}\) sharply, with increase \(W_1\) (see Fig. 21, 22), and, also, although weakly, with increase \(N_2\). Decrease of values \(N_{1\min}^{(T)}\) and \(J_{T\min}\) with increase \(W_1\) is caused by situation when the thicker \(W_1\) the less \(E_{1BD}\) and the greater fall of field \(E(x)\) on \(n_{wg}\) layer thickness. Slight decrease \(N_{1\min}^{(T)}\) and \(J_{T\min}\) with growth \(N_2\) is due to increasing of length of interband tunneling \(l_{Tng}\) in narrow-gap \(n_{ng}\) layer with increase \(N_2\) and the fact that at minimum \(J_{T1} > J_{T2}\). For small values either \(N_2\) or \(W_2\), field \(E(x)\) is so weakly dependent on \(x\) in \(n_{ng}\) layer, that value \(l_{Tng}\) in it is almost constant. Therefore, when \(N_2W_2 < Q(W_1)\) then values \(N_{1\min}^{(T)}\) and \(J_{T\min}\) do no longer depend on \(N_2\) and slightly decrease with increase \(W_2\) due to reducing the length of tunneling generation region in narrow-gap material. In high performance diode, absorber should be punched-through when voltage bias \(V_{b}\) on heterostructure is less than voltage of avalanche breakdown \(V_{BD}\). This eliminates dark diffusion current from narrow-gap layer and increases operational speed. Condition of punch-through of absorber, as follows from 4.1 and 4.2 is given by:

\[
N_1 \times W_1 + N_2 \times W_2 < \frac{E_{1BD}}{q} \times E_{1BD}(N_1, W_1).
\]  

Allowable intervals of concentrations and thicknesses of heterostructure layers are shown in Fig. 21. As can be seen from Fig. 20b, even, when \(R_1 = R_3\) quantum efficiency reaches almost its maximal value when \(W_2 = 2\) μm. Therefore, for development of concentration – thickness nomogram in Fig. 21, namely this value \(W_2\) was selected. Note that decrease in dispersion in \(N_4\) results in increase in dispersion \(N_1\) and \(W_1\), while increase gives the opposite result. Value of noise current density \(I_N \leq 10^{-12}\) A/Hz\(^{1/2}\) corresponds to \(J_T \leq 1.8 \times 10^{-5}\) A/cm\(^2\), and value \(I_N \leq 10^{-13}\) A/Hz\(^{1/2}\) corresponds to \(J_T \leq 1.8 \times 10^{-7}\) A/cm\(^2\).

5.3. Operational speed

Minimal possible time-of-response of this class of devices

\[
\tau = 2 \times (\tau_{tr1} \times f(M_{ph}) + \tau_{tr2})
\]

is determined by time-of-flight of charge carriers through multiplication layer \(\tau_{tr1}\) and absorber \(\tau_{tr2}\), and also by value of function \(f(M_{ph})\), which is close to 1 when \(K >> 1\), and is
Figure 21. Concentration-thickness nomogram for avalanche InP/In$_{0.53}$Ga$_{0.47}$As/InP heterophotodiode when $N_2=(1\pm 5)\times 10^{15}\text{cm}^{-3}$, $W_2=2\ \mu\text{m}$, $M_{ph}=15$, cross-section area $A=5\times 10^3\ \mu\text{m}^2$. When noise current $I_N=\sqrt{S_N}\leq 10^{-13}\ \text{A/Hz}^{1/2}$, then allowable set of points in space $(N_1, W_1)$ lies inside figure a-b-c-d; when $I_N=\sqrt{S_N}\leq 10^{-12}\ \text{A/Hz}^{1/2}$ − inside figure a-e-f-g. Dashed and dash-dot curves − dependences $N_{1\min}(W_1)$ and $\bar{N}_1(W_1)$, respectively: 1 − $N_2=10^{15}\text{cm}^{-3}$, 2 − $N_2=5\times 10^{15}\text{cm}^{-3}$. $N_1$ is measured in units of $10^{16}\text{cm}^{-3}$, $W_1$ − in $\mu\text{m}$. 

http://dx.doi.org/10.5772/50778
equal to $M_{ph}$ when $K = 1$ (Tsang, 1985), (Filachev et al, 2011), (Emmons, 1967), (Kurochkin & Kholodnov, 1996). It was noted above that in $\text{InP} / \text{InxGa}_{1-x}\text{As}$ heterostructure quantum efficiency value $\eta$ lies in interval $0.5 \leq \eta \leq 0.686$ when $R_3 = 1$ and $W_2 \geq 0.5 \mu m$. It means that, because of not so much loss in quantum efficiency $\eta$ compared to maximal possible (only 27% less), time-of-response value $\tau_{tr2} = 5$ ps can be achieved by forming absorber with thickness

$$
\tau \geq 2 \times \left( \tau_{tr1} \times M_{ph} + \tau_{tr2} \right).
$$

As is evident from Fig. 20b, in $\text{InP} / \text{InxGa}_{1-x}\text{As}/\text{InP}$ heterostructure quantum efficiency value $\eta$ lies in interval $0.5 \leq \eta \leq 0.686$ when $R_3 = 1$ and $W_2 \geq 0.5 \mu m$. It means that, because of not so much loss in quantum efficiency $\eta$ compared to maximal possible (only 27% less), time-of-response value $\tau_{tr2} = 5$ ps can be achieved by forming absorber with thickness
\( W_2 = 0.5 \mu m \) and fully reflecting backside surface. Minimal value \( \tau_{tr1} \) is determined by maximum allowable minimal value \( W_{1 \text{min}} \). When \( J_T \leq 10^{-6} \text{ A/cm}^2 \), then as follows from Fig. 22, \( W_{1 \text{min}} \approx 2 \mu m \), and therefore \( \tau_{\text{min}} \approx (4M_{ph} + 1) \times 10^{-2} \text{ ns} \).

6. Analytical model of avalanche photodiodes operation in Geiger mode

We consider possibility to describe transient phenomena in \( p-i-n \) APDs by elementary functions, first of all, when initially applied voltage \( V_0 \) is greater than avalanche breakdown voltage \( V_{BD} \). Formulation of the problem is caused by need to know specific conditions of APDs operation in Geiger mode. Simple expression describing dynamics of avalanche Geiger process is derived. Formula for total time of Geiger process is obtained. Explicit analytical expression for realization of Geiger mode is presented. Applicability of obtained results is defined. APDs in Geiger mode (pulsed photoelectric signals) make possible detection of single photons (Groves et al, 2005), (Spinelli & Lacaita, 1997), (Zheleznykh et al, 2011), (Stoppa et al, 2005), (Gulakov et al, 2007). It is worked at reverse bias voltages \( V_b > V_{BD} \). Different types of devices are realized on APDs in Geiger mode (Groves et al, 2005), (Spinelli & Lacaita, 1997), (Zheleznykh et al, 2011), (Stoppa et al, 2005), (Gulakov et al, 2007). At the same time, review of publications shows that theoretical studies have tendency to carry out increasingly sophisticated numerical simulations. In (Vanyushin et al, 2007) was proposed discrete model of Geiger avalanche process in \( p-i-n \) structure. Obtained iterative relations allow to determine, although fairly easy, but only by numerical method, options for realization of Geiger mode when ratio \( K \equiv \beta / \alpha \) differs very much from unity, where \( \alpha(E_{i \text{GE}}) \) and \( \beta(E_{i \text{GE}}) \) – impact ionization coefficients of electrons and holes and \( E_{i \text{GE}} \) – electric field in \( i \)-layer (base \( 0 < x < W_{i \text{GE}} \), Fig. 23). "Continuous" model (Kholodnov, 2009) developed in this section admits value \( K = 1 \). Considered below approach allows also to describe conditions of realization of Geiger mode and its characteristics by mathematically simple, graphically illustrative relations. It is adopted that photogeneration (PhG) is uniform over sample cross-section area \( S \) transverse to axis \( x \) (Fig. 23). Then, in the most important single-photon process, area \( S \), according to uncertainty principle, shall not exceed in the order of magnitude, square of wavelength of light \( \lambda \). Under these conditions, it is allowably to consider problem as one-dimensional (axis \( x \), Fig. 23). There are grounds to suppose that go beyond one-dimensional model at local illumination make no sense. Single-photon case arises itself when \( S >> S_1 = \pi \times \lambda^2 \). The matter is that charge, during Geiger avalanche process, as show estimates below, has no time to spread significantly over cross section area. Consider serial circuit: \( p-i-n \) diode – load resistance \( R \) – power supply source providing bias \( V_b > V_{BD} \). Let \( p \) and \( n \) regions are heavily doped, so that prevailing share of bias falls across base \( i \). Then after charging process voltage on it can be considered equals to \( V_i = V_0 \). When electron-hole pairs appear in the base then occurs their multiplication that results in decrease \( V_i \), due to screening of field \( E_{i \text{GE}} \) in base by major charge carriers inflowing into \( p \) and \( n \) regions (Fig.
23) in quantity $N_n$ and $P_p$ and voltage drop across load resistor $V_R$ and, hence, current in external circuit arise

$$I_R = \frac{V_R}{R} = \frac{V_b - V_i}{R}. \quad (133)$$

**Figure 23.** Avalanche process in $p-i-n$ structure: 

- “−” – acceptors charge in boundary $i-p$ layer (cathode plate – Cathode);
- “+” – donors charge in boundary $i-n$ layer (anode plate – Anode);
- $\oplus$ and $\ominus$ – generated in $i$ – region avalanche photoelectrons and photoholes;
- $N_n$ and $P_p$ – inflowing in $n$ – and $p$ – regions avalanche photoelectrons and photoholes; $E_c$ and $E_v$ – energy of conduction band bottom and valence band top; $h\nu$ – photon energy

In present structure charge is mainly concentrated in thin near border $n-i$ and $p-i$ layers (let’s call them plates, Fig. 23). Therefore, as in (Vanyushin et al, 2007), field $E_i^{GE}$ will be as-
Numerical value \( E_{GE} = E_{BD} \) when \( V_i = V_{BD} \) for a number of materials can be quickly determined by formulas given in Section 3. As in (Vanyushin et al, 2007), we restrict consideration by PhG in base only, we neglect recombination in it, and we assume that currents of electrons \( I_N \) and holes \( I_P \) are determined by their drift in electric field with velocity of saturation \( v_s \), i.e.,

\[
I_N(x, t) = q \times v_s \times N(x, t), I_P(x, t) = I(x, t) - I_N(x, t) = q \times v_s \times P(x, t),
\]

where \( N \) and \( P \) – linear density (per unit length) of electrons and holes, \( I \) – full conductive current, \( q \) – absolute value of electron charge, \( t \) - time.

Substituting volume charge density from Poisson equation in continuity equation for \( I \) and integrating over depletion layer (DL) we obtain that, in approximation of zero-bias current, in quasi-neutral parts of structure

\[
I_R = C_d \times \frac{\partial V_d}{\partial t} + <I_d>, <I_d> = \frac{1}{W_i^{GE}} \times \int_0^{W_d} I(x, t) dx
\]

where \( V_d \) – voltage on DL, \( C_d = \varepsilon \varepsilon_0 S / W_d \) and \( W_d \) – DL capacity and thickness, \( \varepsilon_0 \) – dielectric constant of vacuum, \( \varepsilon \) – dielectric permittivity, \(<I_d>\) let’s call avalanche current \( I_{\text{av}} \).

Relation (135) generalizes well-known theorem of Rameau (Spinelli & Lacaita, 1997), it takes into account key feature of Geiger mode – variation over time of voltage across DL, and it is valid for any distribution profile of dopant. In our formulation of the problem (in \( p-i-n \) structure) \( i \) – layer can be considered as DL, i.e., \( d \) in (135) and below should be replaced by \( i \). By integrating continuity equation for \( I_N \) and \( I_P \) with respect to \( x \) from 0 to \( W_i^{GE} \) and marking linear density of photogeneration rate as \( G(x, t) \) we obtain equations

\[
q \times \frac{\partial N_i(t)}{\partial t} = \alpha \times \tilde{I}_N(t) + \beta \times \tilde{I}_P(t) + I_N(W_i^{GE}, t) - I_N(0, t) + q \times \tilde{G}(t),
\]

\[
q \times \frac{\partial P_i(t)}{\partial t} = \alpha \times \tilde{I}_N(t) + \beta \times \tilde{I}_P(t) - I_P(W_i^{GE}, t) + I_P(0, t) + q \times \tilde{G}(t),
\]

\[
N_i = \int_0^{W_i^{GE}} N(x, t) dx, P_i = \int_0^{W_i^{GE}} P(x, t) dx, \tilde{I}_N(t) = \int_0^{W_i^{GE}} I_N(x, t) dx, \tilde{G}(t) = \int_0^{W_i^{GE}} G(x, t) dx
\]
Because plates are very thin, then generation and recombination in them can be neglected. Now by integrating same equations with respect to thickness of plates, we find that in approximation of absence of minority carriers in \( p \) and \( n \) regions

\[
I_N(0, t) = I_R + q \times \frac{\partial N_n}{\partial t} = I_R - C_i \times \frac{\partial V_i}{\partial t} = I_P(W_i^{GE}, t), \quad I_P(0, t) = I_N(W_i^{GE}, t) = 0 \tag{139}
\]

Strictly speaking, equations (139) are valid when \( r_1 = P_p/N_n = 1 \), from which \( r_2 = |P_i - N_i| / N_n = 0 \). Therefore, let’s assume uniform PhG along \( x \). Then, at \( K = 1 \), symmetry requires \( r_1 = 1 \). Equations (139) are correct in concern of the order of magnitude both when \( K \) is not too big and when small also. This follows from quasi-discrete computer iterations in uniform static field. Computer iterations are performed in several evenly spaced points of PhG \( x_g \) succeeded by averaging with respect to \( x_g \) and take into account much more number acts of impact ionization by holes than similar iterations in (Vanyushin et al, 2007). Iteration procedure performed in interval equals to several time-of-flight of charge carriers through base \( t_{tr} \) gives 0.6 < \( r_1 \) < 1, and \( r_2 \) < 0.4 (Fig. 24a), which corresponds to approximation of uniform field. Note that smallness \( r_2 \) does not mean smallness \( P_i + N_i \) (curve 3 in Fig. 24a).

Relations (133)-(139) allow obtaining equations

\[
F[V_R; (1 / \tau_j)] = \frac{\partial^2 V_R}{\partial t^2} + \left[ \frac{1}{\tau_j} - v_x \times Y(E_i^{GE}(V_R)) \right] \frac{\partial V_R}{\partial t} + \frac{v_x}{\tau_j} \times Y(E_i^{GE}(V_R)) \times V_R = q \times \overline{G}(t) \times \frac{2 \times v_x}{C_i \times W_i^{GE}}, \tag{140}
\]

with initial conditions

\[
V_R(0) = 0, \quad \left. \frac{\partial V_R}{\partial t} \right|_{t=0} = \frac{2v_x}{C_i \times W_i^{GE}} \times \lim_{t \to 0} \int_{-t}^{t} \overline{G}(t') dt', \tag{141}
\]

where

\[
Y(E_i^{GE}) = X(E_i^{GE}) - (2 / W_i^{GE}), X = \alpha(E_i^{GE}) + \beta(E_i^{GE}), E_i^{GE} = (V_p - V_R) / W_i^{GE}, \tau_i = RC_i, C_i = \varepsilon \varepsilon_0 \times S / W_i^{GE} \tag{142}
\]

At delta-shaped time-evolving illumination \( \overline{G}_i(t) = N_{ph} \times \delta(t) \) relations (140) and (141) are converted into

\[
F[V_R; (1 / \tau_j)] = 0, V_R(0) = 0, \left. \frac{\partial V_R}{\partial t} \right|_{t=0} = A^{GE} = \frac{q \times 2v_x \times N_{ph}}{\varepsilon \varepsilon_0 \times S} \tag{143}
\]

where \( N_{ph} \) – number of absorbed photons.
Figure 24. Evaluation of applicability of quasi-uniform field approximation. (a) – Results of quasi-discrete computer iterative procedure $r_i(K)$: $r_1 = \frac{P_i}{N_n}$, $r_2 = \frac{|P_i - N_i|}{|N_n|}$, $r_3 = \frac{(P_i + N_i)}{(P_i + N_i)}$. (b) – Dependence of error ER during determination of breakdown field on $K = \beta / \alpha$; accepted (Tsang, 1985), (Grekhov & Serezhkin, 1980) $\alpha(E) = A^\text{GE} \times \exp(-B / E)$, where $A^\text{GE}$, $1/\mu$m: 1 – 200, 2 – 400, 3 – 800, 4 – 2000, 5 – 5000
If we take $R=0$ and $\lim_{t \to \infty} G_i(t) = \text{const} \neq 0$ then we find that breakdown is determined by condition $W_i^{GE} \times X(E_i^{GE}) = 2$, which at $K \neq 1$ gives another value for breakdown field $E_i^{GE} = E_{av}$ than $E_i^{GE} = E_{BD}$ obtained directly from solving of stationary problem in Section 3. However, discrepancy between $E_{av}$ and $E_{BD}$ is no more than 20%, if $K$ is different from 1 by no more than two orders of magnitude (Fig. 24b). Equation (140) admits only numerical solution. However, Geiger mode can be described without solving this equation, by using physical grounds and limit $R \to \infty$, when

$$I_{av} = C_i \times \frac{\partial \Delta V_i}{\partial t} F[\Delta V; 0] = 0, E_i^{GE} = V_i / W_i^{GE} = [V_b - \Delta V_i(t)] / W_i^{GE}, \Delta V_i(0) = 0, \frac{\partial \Delta V_i}{\partial t} \big|_{t=0} = A^{GE}$$

and problem is solved in quadratures. To solve in elementary functions let’s approximate exact dependence $Y[E_i^{GE}(\Delta V_i)]$ by piecewise-linear function passing through principal point $\Delta V = D_{av} = V_b - V_{av} = V_b - E_{av} \times W_i^{GE}$ (Fig. 25 and 26), where $Y=0$, and $I_{av}$ reaches its peak during $t_{av}$.

Figure 25. Form of approximation of function $Y(\Delta V)$. Dependences (1 - exact, 2 - approximate) are plotted for Ge with orientation <100> (Tsang, 1985) taken $W_i^{GE} = 1 \mu m, D_{av} = 4 V$.
Figure 26. Ratio of approximate dependence $\tilde{Y}(\Delta V)$ to exact $Y(\Delta V)$ for Ge with orientation $<100>$; $\delta = \Delta V / (\Delta V)_{\text{max}}$ (--- $W_i^{\text{GE}}=1 \, \mu\text{m}$, ( ) $W_i^{\text{GE}}=2 \, \mu\text{m}$, $D_{\text{av}}, V$: 1 − 0.25, 2 − 0.5; 3 − 1).

Suppose, for simplicity $X(E_b) \leq 4 / W_i^{\text{GE}}$, where $E_b = V_b / W_i^{\text{GE}}$. Then $\Delta V_{\text{imax}} = \lim_{t \to \infty} V_i(t)$ is not more than value of break point $\Delta V_k$ of piecewise-linear approximation (Fig. 25). Under these conditions

$$\Delta V_i(t) = \Delta V_{\text{imax}} \times \frac{Z^{1/t_{\text{av}}} - 1}{Z^{1/t_{\text{av}}} + Z} \Delta V_{\text{imax}} = 2D_{\text{av}} t_{\text{av}} \frac{\ln Z}{Z} = \frac{\varepsilon \epsilon_0 \times S \times Y_b \times D_{\text{av}}}{q \times N_{ph}^0} \gg 40,$$

(145)

where $Y_b = Y_b(E_b)$. Geiger mode occurs when during time $R \times C_i$ of inverse recharge of avalanche diode, avalanche is able to develop and cancel itself in full. As seen from (145) it is happened when $R \geq R_{\text{min}} \approx t_{\text{av}} / C_i$. Maximal voltage drop on load equals to $V_{R_{\text{max}}} = \Delta V_{\text{imax}}$. Since $t_{\text{tr}} \ll t_{\text{av}}$, then results of computer evaluation of uniform field approximation applicability can be considered reasonable. To evaluate transverse charge spreading let’s use expres-
sion (21) from (Pospelov et al, 1974). It determines dependence \( \chi(t) \equiv r / r_0 \), where \( r(t) \) and \( r_0 \) – current and initial radii of charge “drop” of parabolic type. Implied under capacity in (Pospelov et al, 1974) value \( C_i \) and putting \( W_i^{GE} = 1 \ \mu m, \ r_0 = \lambda = 1 \ \mu m, \) in the case of single-photon process we get \( \chi(t_{av}) < 2^{1/4} \approx 1.2 \). This justifies our assumption that charge spreading over sample cross-section during avalanche Geiger process is not intensive.

7. Conclusions

The above analysis shows that to create high performance SAM-APD (in particular, based on widely used \( \text{InP} / \text{In}_x\text{Ga}_{1-x}\text{As}_{1-y} / \text{InP} \) heterostructures) it is necessary to maintain close tolerances on dopants concentration in wide-gap multiplication layer \( I – N_1 \) and in narrow-gap absorption layer \( II – N_2 \) and also on thickness \( W_1 \) of wide-gap multiplication layer (Fig. 1). This is due to strong dependence of interband tunnel current in such heterostructures on \( N_1, N_2 \) and \( W_1 \). Allowable variation intervals of values \( N_1, N_2 \) and \( W_1 \) and, optimal thickness of absorber also, can be determined using results obtained in Sections 4 and 5. Value of minimal possible time-of-response \( \tau_{min} \) depends not only on photocurrent’s gain \( M_{ph} \) but on allowable noise density at preset value of photocurrent’s gain also. The lower noise density, the larger is value \( \tau_{min} \). For example, for heterostructure \( \text{InP} / \text{In}_{0.53}\text{Ga}_{0.47}\text{As} / \text{InP} \) minimal time-of-response equals to \( \tau_{min} \approx 0.6 \ \text{ns}, \) when noise current equals to \( 3.3 \times 10^{-11} \ \text{A/Hz}^{1/2} \) and current responsivity \( 10.3 \ \text{A/W}. \) Analysis shows that operational speed can be slightly increased by means of inhomogeneous doping of wide-gap multiplication layer. To ensure operational speed in picosecond range it is necessary to use as multiplication layer semiconductor layer with low tunnel current and impact ionization coefficients of electrons and holes much different from each other, for example, indirect-gap semiconductor silicon. As has long been known maximal operational speed is achieved by APD if light is absorbed in space-charge region. In this case, as it was shown in Section 6, when bias voltage \( V_b \) exceeds breakdown voltage \( V_{BD} \) of no more than a few volts, then, for \( K = \beta / \alpha \) values lying in interval from a few hundredths to a few tens, elementary relations (145) can be used for approximate description of Geiger mode in \( p-i-n \) APD. Moreover if cross-section area \( S > S_1 = \pi \times \lambda^2 \), then we can expect that in single-photon case under \( S \) in (145) should imply value of order \( S_1 \). This is due to finite size of single-photon spot \( S_1 \) and not intensive spreading of charge during time of avalanche Geiger process \( t_w \) when photogeneration of charge carriers occurs in \( i – \text{region of } p-i-n \) structure depleted by charge carriers. Proposed approach allows describing Geiger mode by elementary functions at voltages higher \( V_b \) as well. Note that equation (140) and physical grounds allow to expect three possible process modes at pulse illumination under \( V_b > V_{BD} \). When \( RC < < t_w \) then generated photocurrent will tend to reach some constant and flow indefinitely (unless, of course, ignore energy losses). When \( RC = t_w \) then generated photocurrent will be of infinitely long oscillatory character. When \( RC > > t_w \), then Geiger mode is realized.
Author details

Viacheslav Kholodnov\textsuperscript{1} and Mikhail Nikitin\textsuperscript{2}

\textsuperscript{1} V.A. Kotelnikov Institute of Radio Engineering and Electronics Russian Academy of Sciences, Moscow, Russia

\textsuperscript{2} Science & Production Association ALPHA, Moscow, Russia

References


[29] Ito M.; Kaneda T.; Nakajima K.; Toyama Y. & Ando H. (1981). Tunneling currents in In0.53Ga0.47As homojunction diodes and design of In0.53Ga0.47As/InP heterostructure avalanche photodiodes. Solid-State Electron., Vol. 24, No. 5, (May 1981) p. 421-424, ISSN 0038-1101


