

Методические указания к расчётному заданию по дисциплине «Твердотельная электроника»

Вольт-амперная характеристика идеального *pn*-перехода

Для идеального *pn*-перехода зависимость тока через него от напряжения, прикладываемого к переходу (вольт-амперная характеристика), описывается уравнением Шокли:

$$I(U, T) = I_S(T) \cdot \left[\exp\left(\frac{U}{\phi_T(T)}\right) - 1 \right], \quad (1)$$

где I_S – ток насыщения, А,

$$\phi_T(T) = \frac{k_B T}{q} \quad \text{– тепловой потенциал, В;}$$

T – температура, К;

$k_B = 1,38 \cdot 10^{-23}$ Дж/К – постоянная Больцмана в Дж/К;

$q = 1,6 \cdot 10^{-19}$ Кл – элементарный заряд.

Ток насыщения I_S является током **неосновных** носителей заряда, которые диффундируют в квазинейтральных областях. Величина плотности тока насыщения ($j_S = I_S / S_{pn}$, где S_{pn} – площадь *pn*-перехода) складывается из плотности тока насыщения электронов j_{ns} , которые диффундируют в *p*-области, и дырок j_{ps} , которые диффундируют в *n*-области:

$$j_S(T) = j_{ns}(T) + j_{ps}(T) = q \frac{D_n(T)}{L_n(T)} n_{p0}(T) + q \frac{D_p(T)}{L_p(T)} p_{n0}(T), \quad (2)$$

где n_{p0} , p_{n0} – равновесные концентрации **неосновных** носителей заряда (электронов в *p*-области и дырок в *n*-области), см⁻³;

D_n , D_p – коэффициенты диффузии **неосновных** носителей заряда, см²/с;

L_n , L_p – диффузионные длины **неосновных** носителей заряда, см.

Для определения концентрации **неосновных** носителей заряда используется связь с концентрацией основных носителей заряда:

$$\begin{cases} n_{n0}(T) \cdot p_{n0}(T) = n_i^2(T), \\ p_{p0}(T) \cdot n_{p0}(T) = n_i^2(T), \end{cases} \quad (3)$$

где $n_{n0}(T) \equiv N_d$, $p_{p0}(T) \equiv N_a$ – равновесная концентрация основных носителей заряда равная степени легирования соответствующей области, см^{-3} ;

n_i – собственная концентрация, см^{-3} .

Для расчёта собственной концентрации, которая является параметром полупроводникового материала, необходимо знать ширину запрещённой зоны $E_g(T)$, а также эффективную плотность квантовых состояний в зоне проводимости $N_c(T)$ и валентной зоне $N_v(T)$:

$$n_i(T) = \sqrt{N_c(T)N_v(T)} \exp\left[-\frac{E_g(T)}{2k_B T}\right], \quad (4)$$

где $N_c(T) = N_{c0} \left(\frac{T}{300}\right)^{3/2}$;

$$N_v(T) = N_{v0} \left(\frac{T}{300}\right)^{3/2} ;$$

N_{c0} , N_{v0} – эффективные плотности квантовых состояний в зоне проводимости и валентной зоне при температуре 300 К, см^{-3} ;

$E_g(T) = E_g^0 + \alpha \cdot T$ – ширина запрещённой зоны, эВ;

E_g^0 – ширина запрещённой зоны при температуре 0 К, эВ;

α – температурный коэффициент ширины запрещённой зоны, эВ/К;

здесь $k_B = 8,6 \cdot 10^{-5}$ эВ/К – постоянная Больцмана в эВ/К.

Коэффициент диффузии **неосновного** носителя заряда определяется из величины его подвижности согласно соотношению Эйнштейна:

$$\begin{cases} D_n(T) = \varphi_T(T) \cdot \mu_n(N_a, T), \\ D_p(T) = \varphi_T(T) \cdot \mu_p(N_d, T), \end{cases} \quad (5)$$

где μ_n , μ_p – подвижности **неосновных** носителей заряда, $\text{см}^2/(\text{В} \cdot \text{с})$.

Величина подвижности определяется не только температурой структуры, но также и концентрацией примеси, на которой рассеивается носитель заряда. Так, в случае **неосновных** носителей заряда, если рассматривается движение электрона в p -области, то рассеиваться он может только на ионах акцепторов N_a , а дырка, двигающаяся в n -области, может рассеиваться только на ионах доноров N_d . Поэтому для **неосновных** носителей заряда $\mu_n \equiv \mu_n(N_a, T)$ и $\mu_p \equiv \mu_p(N_d, T)$.

Диффузионная длина **неосновных** носителей заряда связана со временем их жизни:

$$\begin{cases} L_n(T) = \sqrt{D_n(T) \cdot \tau_n}, \\ L_p(T) = \sqrt{D_p(T) \cdot \tau_p}, \end{cases} \quad (6)$$

где τ_n, τ_p – времена жизни **неосновных** носителей заряда (электронов и дырок соответственно), с.

Построенная на основе рассчитанных величин ВАХ идеального pn -перехода, определяемая уравнением (1), представлена на рисунке 1. На нём отмечено, что при прямом смещении ($U > 0$) ток экспоненциально возрастает и величина напряжения при $I \rightarrow \infty$ стремится к контактной разности потенциалов

$$\varphi_k(T) = \varphi_T(T) \cdot \ln \left[\frac{N_a \cdot N_d}{n_i^2(T)} \right]. \quad (7)$$

При обратном смещении ($U < 0$) ток ограничен током насыщения $I \rightarrow I_s$.

Вольт-амперная характеристика реального диода

Вольт-амперная характеристика реального диода отличается от характеристики идеального pn -перехода тем, что учитывается влияние *процессов генерации и рекомбинации* в области пространственного заряда, а также *сопротивление слаболегированной квазинейтральной области* – базы.

Процессы генерации и рекомбинации в области пространственного заряда приводят к появлению **генерационно-рекомбинационного тока**

$$I_{GR}(U, T) = \frac{qn_i(T)W_{\text{ОПЗ}}(U, T)}{\tau_{\text{эфф}}} S_{pn}, \quad (8)$$

где

$$W_{\text{ОПЗ}}(U, T) = \sqrt{\frac{2\varepsilon_0\varepsilon_r}{q} [\varphi_K(T) - U] \left(\frac{1}{N_a} + \frac{1}{N_d} \right)} \quad (9)$$

– ширина области пространственного заряда, см;

$\varepsilon_0 = 8,85 \cdot 10^{-14}$ Ф/см – диэлектрическая постоянная;

ε_r – относительная диэлектрическая проницаемость материала полупроводника;

$$\tau_{\text{эфф}} = \sqrt{\tau_n \tau_p} \ln \left(\frac{\tau_p}{\tau_n} \right) \quad (10)$$

– эффективное время жизни, с.

Учёт влияния процессов генерации и рекомбинации в ОПЗ видоизменяет выражение (1) для идеального pn -перехода в следующее

$$I(U, T) = I_S(T) \left[\exp \left(\frac{U}{\varphi_T(T)} \right) - 1 \right] + I_{GR}(U, T) \left[\exp \left(\frac{U}{2 \cdot \varphi_T(T)} \right) - 1 \right]. \quad (11)$$

Наличие сопротивления слаболегированной квазинейтральной области – **базы** – приводит к тому, что внешнее напряжение на диоде U_D будет падать на объёмное сопротивление базы R_B и pn -переход

$$U_D(U_{pn}, T) = U_{pn} + I(U_{pn}, T) \cdot R_B(U_{pn}, T), \quad (12)$$

где U_{pn} – напряжение на pn -переходе, В;

$I(U_{pn}, T)$ – ток через диод, определяемый выражением (11) с заменой U на U_{pn} , А;

$R_B(U_{pn}, T)$ – сопротивление квазинейтральной базовой области, Ом.

Сопротивление базы зависит от величины приложенного к pn -переходу напряжения по той причине, что в результате модуляции ширины области пространственного заряда происходит изменение ширины квазинейтральной области

$$R_B(U_{pn}, T) = \frac{1}{\sigma_B(T)} \cdot \frac{l_B(U_{pn}, T)}{S_{pn}} = \frac{1}{\sigma_B(T)} \cdot \frac{d_B - W_B(U_{pn}, T)}{S_{pn}}, \quad (13)$$

где σ_B – удельная электропроводность базы, Ом⁻¹см⁻¹;

l_B – эффективная ширина базы, см;

d_B – технологическая ширина базы, см;

W_B – уширение области пространственного заряда в базу, см.

При расчёте электропроводности базовой области следует помнить, что она определяется **основными** носителями заряда, соответственно в базовой n -области, электрон, будучи основным носителем заряда рассеивается на ионах доноров N_d , а в базовой p -области дырка, будучи основным носителем заряда, рассеивается на ионах акцепторов N_a . Поэтому величины подвижности носителей заряда будут **отличаться** от рассчитанных для **неосновных** носителей заряда при расчёте тока насыщения. Для **основных** носителей заряда $\mu_n \equiv \mu_n(N_d, T)$ и $\mu_p \equiv \mu_p(N_a, T)$.

Уширение области пространственного заряда для n - и p -базы определяется соответственно следующим образом:

$$\begin{cases} W_n(U_{pn}, T) = W_{\text{ОПЗ}}(U_{pn}, T) \cdot \frac{N_a}{N_a + N_d}, \\ W_p(U_{pn}, T) = W_{\text{ОПЗ}}(U_{pn}, T) \cdot \frac{N_d}{N_a + N_d}. \end{cases} \quad (14)$$

Аналогично рассчитывается сопротивление сильнолегированной области – **эмиттера**. После расчёта можно убедиться, что благодаря большой концентрации легирующей примеси и малой ширине области сопротивление эмиттера существенно мало по сравнению с сопротивлением базы $R_E \ll R_B$, и как следствие на нём не учитывается падение внешнего напряжения.

Используя выражения (11) и (12), мы получаем систему уравнений, описывающих ВАХ **реального** диода $I(U_D)$

$$\begin{cases} I(U_{pn}, T) = I_S(T) \left[\exp\left(\frac{U_{pn}}{\Phi_T(T)}\right) - 1 \right] + I_{GR}(U_{pn}, T) \left[\exp\left(\frac{U_{pn}}{2 \cdot \Phi_T(T)}\right) - 1 \right], \\ U_D(U_{pn}, T) = U_{pn} + I(U_{pn}, T) \cdot R_B(U_{pn}, T), \end{cases} \quad (15)$$

в которой U_{pn} является по сути параметром.

При построении этой характеристики в программе MathCad (или Matlab-подобной) необходимо сначала задать значение параметра U_{pn} в диапазоне, например, от -2 до +2 В с шагом 1 мВ, а затем построить зависимость $I(U_{pn})$ от $U_D(U_{pn})$.

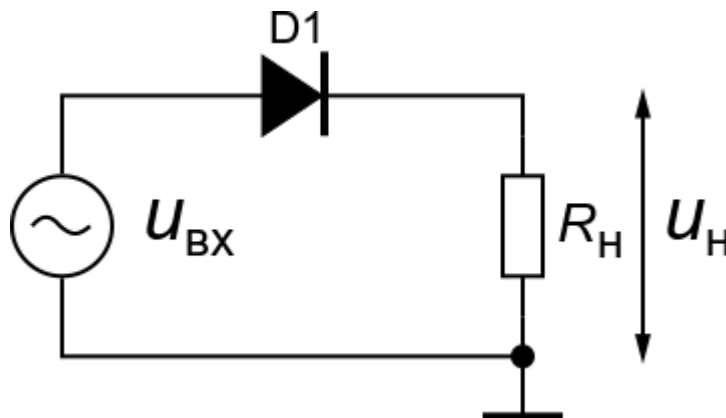
С увеличением обратного смещения возрастает вероятность пробоя, один из видов которого – *лавинный*. Напряжение, при котором наступает этот пробой можно оценить по эмпирической формуле

$$U_{пр} = 60[\text{В}] \cdot \left(\frac{E_g [\text{эВ}]}{1,1[\text{эВ}]} \right)^{3/2} \cdot \left(\frac{N_B [\text{см}^{-3}]}{10^{16} [\text{см}^{-3}]} \right)^{-3/4}, \quad (16)$$

где N_B – концентрация легирующей примеси в слаболегирующей области – базе, см^{-3} .

Выпрямление на диоде

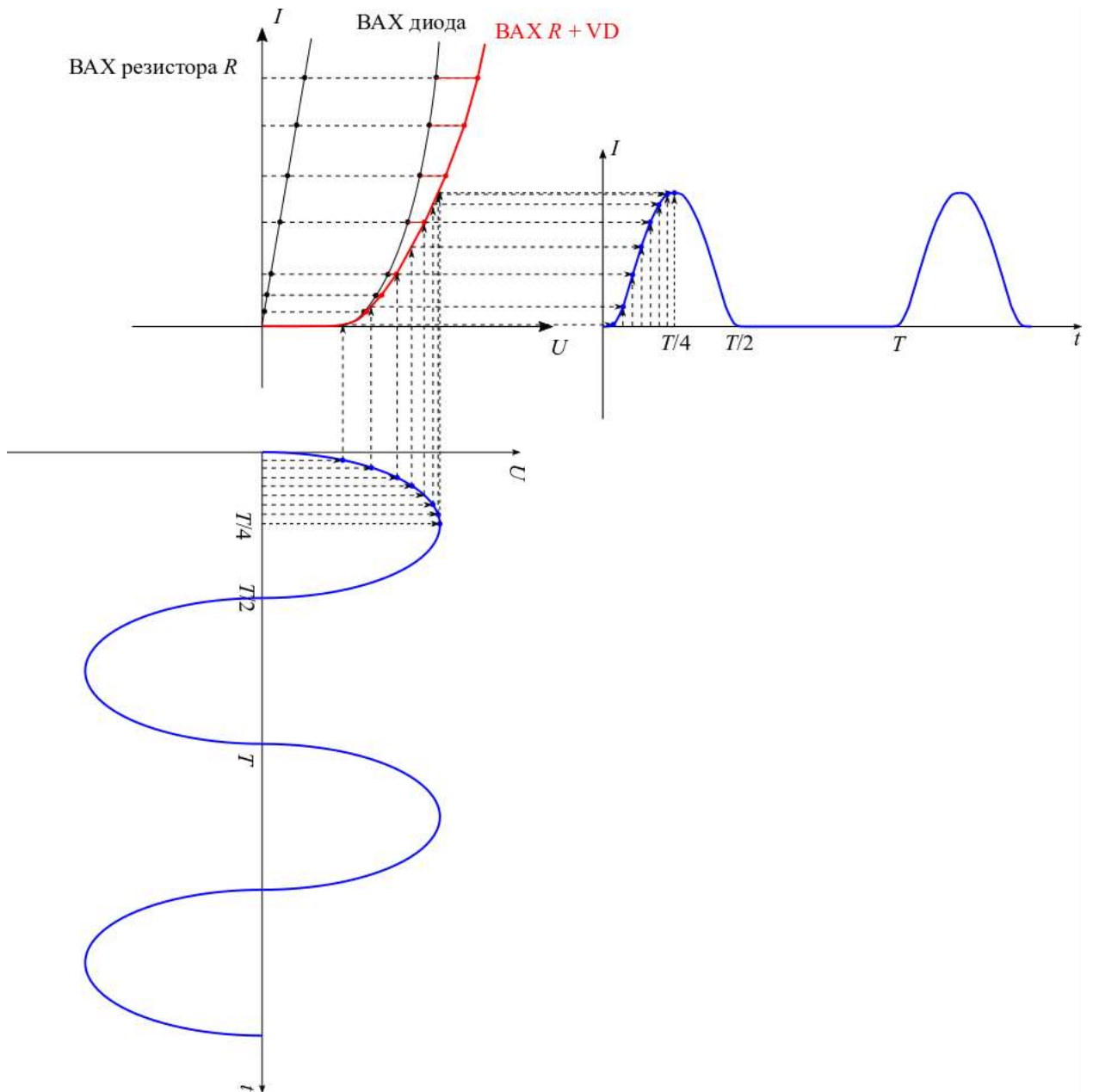
Благодаря своей вольт-амперной характеристике диод обладает выпрямляющими свойствами, что определяется его способностью пропускать ток только в одном направлении – прямом, от анода к катоду. В обратном направлении, от катода к аноду, ток пренебрежимо мал. Это хорошо заметно, если подключить к нагрузочному резистору R_H источник синусоидального напряжения $u_{вх}(t)$ через диод D1.



Напряжение на резисторе будет определяться согласно закону Ома током через него $u_{\text{н}}(t) = R_{\text{н}} \cdot i(t)$. Для определения тока $i(t)$ с учётом наличия нелинейного элемента – диода – необходимо построить общую вольт-амперную характеристику диода и резистора. Так как эти элементы подключены последовательно, то напряжение $u_{\text{вх}}$ есть сумма напряжений на резисторе и диоде, а ток у них одинаковый. Это означает, что вольт-амперные характеристики диода и резистора необходимо сложить по оси напряжения и получить общую вольт-амперную характеристику (красная кривая на графике). Чем больше используется точек в области «пяточки», или изгиба ВАХ, тем точнее получится общая кривая.

Для построения зависимости $i(t)$ на миллиметровой бумаге используется понятие эпюр напряжения и токов. Ось напряжений «раздвигают» во временную область по вертикали, а ось токов – по горизонтали. На эпюре напряжений изображают осциллограмму входного напряжения, а на эпюре тока будет формироваться осциллограмма тока следующим образом. Задаются моментом времени (например, четверть периода $T/4$) на осциллограммах напряжения и тока. На эпюре напряжения определяют величину напряжения в этот момент времени, после чего его проецируют на общую вольт-амперную характеристику схемы, определяя тем самым ток в этот момент времени. Это значение тока переносится на эпюру тока в тот же момент времени $T/4$. Затем задаются другим моментом времени и повторяют операцию с проецированием. Чем больше моментов времени будет использовано, тем точнее получится осциллограмма тока.

Исходя из периодичности входного синусоидального сигнала, а также из того, что диод и резистор не вносят каких-либо изменений в эту периодичность, то частота (период) тока будет таким же, как и у входного напряжения. Поэтому построить методом проецирования достаточно один период, после чего его продублировать.



Ёмкостные свойства диода

Ёмкостные свойства диода характеризуются *барьерной* и *диффузионной* ёмкостями.

Барьерная ёмкость является следствием модуляции ширины ОПЗ при изменении внешнего напряжения. Область пространственного заряда в этом случае представляет собой прослойку в плоском конденсаторе, и барьерная ёмкость определяется как

$$C_{\text{бар}}(U) = \frac{\epsilon_0 \epsilon_r}{W_{\text{ОПЗ}}(U)} S_{pn} = \frac{C_{\text{бар}}^0}{\sqrt{1 - \frac{U}{\Phi_K}}}, \quad (17)$$

где $C_{\text{бар}}^0 = \frac{\varepsilon_0 \varepsilon_r}{W_{\text{ОПЗ}}^0} S_{pn}$ – барьерная ёмкость в отсутствие напряжения, Ф;

$W_{\text{ОПЗ}}^0$ – ширина ОПЗ в отсутствие напряжения, см.

Анализируя выражение (17), можно заметить, что зависимость $C_{\text{бар}}^{-2}(U)$ линейна. В таких координатах линия отсекает на оси напряжений величину контактной разности потенциалов, а её наклон позволяет экспериментально оценить степень легирования одной из областей диода (при известной другой).

Диффузионная ёмкость возникает в результате инжекции неосновных носителей заряда и определяет зависимость количества инжектированных носителей заряда при модуляции напряжения

$$C_{\text{диф}}(U) = \frac{dQ_{\text{инж}}}{dU} \approx \frac{I(U) \cdot \tau_{\text{эфф}}}{\varphi_T}, \quad (18)$$

где $Q_{\text{инж}}$ – инжектированный заряд, Кл;

$I(U)$ определяется выражением (11).

Инжектированный заряд определяется распределением концентрации неосновных носителей заряда в квазинейтральных областях. В случае только диффузионного движения распределение будет зависеть от ширины области l по сравнению с величиной диффузионной длины неосновного носителя заряда L . Разделяют *длинную* ($l \gg (3-5)L$) и *короткую* ($l < (3-5)L$) области.

В случае длинной области распределение неосновных носителей заряда характеризуется убывающей экспонентой

$$\begin{cases} n_p(x, U) = n_{p0} + \Delta n_p(U) \cdot \exp\left(-\frac{x}{L_n}\right), \\ p_n(x, U) = p_{n0} + \Delta p_n(U) \cdot \exp\left(-\frac{x}{L_p}\right), \end{cases} \quad (19)$$

где $\Delta n_p(U) = n_{p0} \cdot \exp\left(\frac{U}{\Phi_T}\right) - n_{p0}$ – неравновесная концентрация неосновных

носителей заряда в p -области при напряжении U , см⁻³;

$\Delta p_n(U) = p_{n0} \cdot \exp\left(\frac{U}{\Phi_T}\right) - p_{n0}$ – неравновесная концентрация

неосновных носителей заряда в n -области при напряжении U , см⁻³.

В случае короткой области распределение имеет более резкий спад¹

$$\left\{ \begin{array}{l} n_p(x, U) = n_{p0} + \Delta n_p(U) \cdot \frac{\text{sh}\left(\frac{l_p - x}{L_n}\right)}{\text{sh}\left(\frac{l_p}{L_n}\right)}, \\ p_n(x, U) = p_{n0} + \Delta p_n(U) \cdot \frac{\text{sh}\left(\frac{l_n - x}{L_p}\right)}{\text{sh}\left(\frac{l_n}{L_p}\right)}, \end{array} \right. \quad (20)$$

где l_p, l_n – эффективные ширины p - и n -областей соответственно, см.

Для удобства построения графиков распределения концентрации неосновных носителей заряда в выражении (20) уширением ОПЗ по сравнению с толщиной области можно пренебречь (таким образом $l_n \rightarrow d_n$ и $l_p \rightarrow d_p$).

С точки зрения более строгого определения типа области (длинная или короткая) имеет смысл сначала определить при данном напряжении U_0 концентрацию неосновных носителей заряда в месте металлического контакта (на границе области, например $n_p(d_p, U_0)$) по выражению (19), если она с точностью 1 % совпадает с равновесной концентрацией неосновных носителей заряда n_{p0} , то область считается длинной. Если эта концентрация больше $n_p(d_p, U_0) > n_{p0}$, то короткой.

¹ В программном пакете MathCad функция sh записывается как sinh.

Статическое и дифференциальное сопротивления

Для характеристики работы нелинейных элементов используют понятия *статического* и *динамического* сопротивления.

Статическое сопротивление в точке (U_0, I_0) на вольт-амперной характеристике определяется по закону Ома как

$$R_{\text{ст}} = \frac{U_0}{I_0}. \quad (21)$$

При воздействии на нелинейный элемент малым сигналом, например, синусоидальным напряжением $u(t) = u_m \sin(\omega t)$, амплитуда u_m которого существенно меньше величины постоянного сигнала $u_m \ll U_0$, реакция на этот сигнал будет определяться **динамическим** (или **дифференциальным**) сопротивлением, которое определяется как

$$r_{\text{диф}} = \left. \frac{dU_{\text{Д}}}{dI} \right|_{(U_0, I_0)} \approx \left. \frac{\Delta U_{\text{Д}}}{\Delta I} \right|_{(U_0, I_0)}, \quad (22)$$

где $\Delta U_{\text{Д}}$, ΔI – малые приращения напряжения и тока (линеаризация нелинейной характеристики), которые определяются графически.

Динамическое сопротивление зависит от статического состояния нелинейного элемента, которое характеризуется *рабочей точкой* (U_0, I_0) . По своей сути, это сопротивление определяет наклон касательной в этой точке. Ток (и амплитуда тока) при воздействии малого сигнала в виде синусоидального напряжения будет определяться как $i(t) = u(t) / r_{\text{диф}} \Rightarrow i_m = u_m / r_{\text{диф}}$.

Если к диоду приложить переменное напряжение, имеющее большую постоянную и малую переменную составляющие $U(t) = U_0 + u(t)$, то ток через него $I(t) = I_0 + i(t)$ будет определяться статическим сопротивлением (для постоянной составляющей) и дифференциальным (для малого сигнала):

$$\begin{cases} I_0 = U_0 / R_{\text{ст}}, \\ i(t) = u(t) / r_{\text{диф}}. \end{cases} \quad (23)$$

