К ТЕОРИИ РЕКОМБИНАЦИИ ХОЛЛА–ШОКЛИ–РИДА

© В.А.Холоднов

Государственный научный центр Российской Федерации НПО «Орион»,
111123 Москва, Россия
(Получена 26 июня 1995 г. Принята к печати 10 июля 1995 г.)

Детально проанализировано, как влияет значение концентрации рекомбинационных центров N на процесс генерации–рекомбинации Холла–Шокли–Рида. Рассмотрено слабое отклонение полупроводника от состояния термодинамического равновесия. Показано, что в концентрационных зависимостях времен жизни неравновесных электронов $\tau_n(N)$ и дырок $\tau_p(N)$, в целом падающих с увеличением N, при определенных условиях может быть и участок роста на несколько порядков. Проведено аналитическое решение задачи на экстремум функций $\tau_n(N)$ и $\tau_p(N)$. Исследованы зависимости положений экстремумов и значений этих функций в экстремальных точках от концентрации мелкой легирующей примеси N_D, энергии рекомбинационного уровня $\mathcal{E}_t$, отношения вероятностей захвата дырки и электрона на него θ, энергии запрещенной зоны полупроводника $\mathcal{E}_g$, а также от температуры T. Дано подробное физическое толкование полученных результатов.

1. Введение

Рекомбинация неравновесных электронов и дырок в полупроводниках через глубокие примесные центры во многих случаях является доминирующей [1,2]. Основы теории примесной рекомбинации носителей были заложены еще в 1952 г. в вышедших почти одновременно работах Холла [3] и Шокли и Рида [4]. В них был рассмотрен генерационно-рекомбинационный процесс при одиночном примесном уровне в запрещенной зоне полупроводника. В дальнейших многочисленных работах теория примесной рекомбинации носителей развивалась в различных аспектах и детализировалась [2,5,6]. При этом должного внимания не уделялось исследованию зависимости времен жизни неравновесных электронов $\tau_n(N)$ и дырок $\tau_p(N)$ от концентрации рекомбинационных центров N, которые в ряде случаев преднамеренно вводятся в полупроводник (например, за счет облучения ионами высоких энергий [7,8]) для уменьшения времени переходных процессов. Видимо, это обусловлено сложившимся представлением, что чем больше N, тем больше темп захвата неравновесных носителей на примесные центры, а поэтому тем меньше времена жизни носителей. Такое явно разумное представление не вполне адекватно действительности —

времена жизни неравновесных носителей могут и сильно вырасти (на порядок и более) при увеличении N, как это показано в кратком сообщении [9] на примере времени жизни дырок. Однако работа [9] носит всего лишь оценочный характер. В данной работе проведен последовательный математический и подробный физический анализ зависимости времени жизни неравновесных носителей от концентрации рекомбинационных центров. При этом в отличие от работы [9] исследована не только зависимость $\tau_p(N)$, но и $\tau_n(N)$. Рассмотрено слабое отклонение концентраций свободных носителей от их равновесных значений. Такая ситуация часто реализуется в полупроводниках, когда они используются для регистрации слабых сигналов, например, оптических. Показано, что несмотря на рост темпа захвата неравновесных электронов и дырок с ростом N зависимости как $\tau_p(N)$, так и $\tau_n(N)$ при определенных условиях будут сильно немонотонными. Проанализированы их экстремальные точки (выведены соответствующие формулы) как функции параметров полупроводника и температуры. Проведена детальная физическая интерпретация полученных результатов. В частности, показано, что физические механизмы, ответственные за немонотонность $\tau_p(N)$ и $\tau_n(N)$, отличаются друг от друга.

2. Постановка задачи. Времена жизни неравновесных носителей

Рассмотрим полупроводник, легированный мелкой, полностью ионизованной примесью одного типа (для определенности донорной) с концентрацией N_D, в котором рекомбинация неравновесных (избыточных) носителей происходит через атомы акцепторных примесей, находящиеся в двух зарядовых состояниях (пусть в нейтральном и однократно отрицательно заряженном). Нейтральному состоянию рекомбинационной примеси соответствует концентрация атомов N_0, являющихся центрами захвата электронов и в то же время центрами тепловой генерации дырок, а заряженному — концентрация атомов $N_- = N - N_0$, которые являются центрами захвата дырок и вместе с тем центрами тепловой генерации электронов. Такая ситуация отвечает одиночному рекомбинационному уровню [3,4,10], который во многих случаях является доминирующим [2,11–13].

Пусть неравновесные электроны и дырки возникают либо за счет межзонной генерации, либо за счет инжекционно-контактных процессов. Тогда в стационарном случае зарядовое состояние атомов рекомбинационной примеси определяется уравнением

$$R_n = R_p, \tag{1}$$

где скорости рекомбинации–генерации электронов R_n и дырок R_p, обусловленные захватом носителей акцепторной примесью и тепловым выбросом их с рекомбинационного уровня в разрешенные зоны, равны

$$R_n = nN_0w_n - \delta^{-1}n_eN_-w_n, \qquad R_p = pN_-w_p - \delta p_eN_0w_p. \tag{2}$$

Здесь n и p — концентрации электронов и дырок, n_e и p_e — их равновесные значения, w_n и w_p — вероятности захвата электрона и дырки на соответствующий центр, $\delta = N_-^e/N_0^e$, верхний индекс «e» указывает на равновесные значения концентраций рекомбинационных центров в соответствующих зарядовых состояниях.

При малых отклонениях концентраций носителей $\Delta n = n - n_e$, $\Delta p = p - p_e$ и центров их захвата $\Delta N_0 = N_0 - N_0^e = -\Delta N_- = N_-^e - N_-$ от своих равновесных значений можно произвести линеаризацию соотношений (1) и (2) относительно этих отклонений. Тогда, учитывая уравнение Пуассона

$$\mathrm{div}\Delta E = \frac{4\pi q}{\varepsilon}(\Delta p - \Delta n - \Delta N_-), \tag{3}$$

получим, что

$$R_n = \frac{\Delta n}{\tau_n} + a_n \mathrm{div}\Delta E, \qquad R_p = \frac{\Delta p}{\tau_p} + a_p \mathrm{div}\Delta E, \tag{4}$$

где

$$\frac{1}{\tau_n} = w_n N \frac{\delta\theta}{1+\delta} \frac{N + (1+\delta)(1+\delta^{-1})(n_e + p_e)}{\delta\theta N + (1+\delta)(1+\delta^{-1})(n_e + \delta\theta p_e)}, \tag{5}$$

$$\frac{1}{\tau_p} = w_p N \frac{\delta}{1+\delta} \frac{\delta N + (1+\delta)^2(n_e + p_e)}{\delta N + (1+\delta)^2(n_e + \delta\theta p_e)}, \tag{6}$$

$$a_n = \frac{\varepsilon}{4\pi q} w_n N n_e \frac{\theta(1+\delta)}{\delta\theta N + (1+\delta)(1+\delta^{-1})(n_e + \delta\theta p_e)}, \tag{7}$$

$$a_p = -\frac{\varepsilon}{4\pi q} w_p N p_e \frac{1+\delta}{N + (1+\delta)(1+\delta^{-1})(n_e + \delta\theta p_e)}, \tag{8}$$

ΔE — изменение напряженности электрического поля, связанное с отклонением концентраций носителей и их центров захвата от своих равновесных значений; ε — диэлектрическая проницаемость; q — заряд электрона, $\theta = w_p/w_n$. Первые слагаемые в выражениях (4) означают скорости рекомбинации избыточных электронов и дырок (а поэтому τ_n и τ_p означают их времена жизни) в условиях квазинейтральности по отношению к полю ΔE, т. е. при достаточно малых значениях $|\mathrm{div}\Delta E|$ [1–6,11–13]. Эту терминологию для величин τ_n и τ_p мы оставляет и в случае нарушения квазинейтральности, так что значения τ_n и τ_p, которые мы будет исследовать, не зависят от значения $\mathrm{div}\Delta E$.

Используя функцию распределения электронов по акцепторным примесным состояниям и уравнение нейтральности для невырожденного полупроводника в состоянии термодинамического равновесия [14]

$$\frac{N}{N_D} = \frac{1+\delta}{A\delta^2}(B + A\delta - \delta^2), \tag{9}$$

выражения (5) и (6) можно преобразовать к виду

$$\frac{1}{\tau_n} = w_n N_D \theta \frac{B + A\delta - \delta^2}{A\delta^2} \frac{A\delta + B(2+\delta) + \delta^3}{\theta(B + A\delta - \delta^2) + (1+\delta)(\theta B + \delta)}, \tag{10}$$

$$\frac{1}{\tau_p} = w_p N_D \, \frac{B + A\delta - \delta^2}{A\delta} \, \frac{A\delta + B(2+\delta) + \delta^3}{B + (A + \theta B)\delta + (\theta B + \delta)\delta^2}, \tag{11}$$

где

$$A = 2N_D/n_t, \qquad B = 4p_t/n_t, \tag{12}$$

n_t и p_t — равновесные концентрации электронов и дырок, когда энергия уровня Ферми $\mathcal{E}_F$ совпадает с энергией примесного уровня $\mathcal{E}_t$. При выводе выражений (9)–(12) считалось, что фактор спинового вырождения акцепторного состояния равен 1/2 [5,13,15]. Соотношения (9)–(12) в параметрическом виде определяют зависимость времен жизни неравновесных носителей от концентрации рекомбинационных центров. Из рис. 1 видно, что в зависимостях $\tau_n(N)$ и $\tau_p(N)$, в целом падающих с увеличением N, может быть и участок резкого роста.

3. Экстремальные точки

Далее проведено аналитическое решение задачи на экстремумы функций $\tau_n(N)$ и $\tau_p(N)$ при $\theta \geqslant 1$, так как дырки захватываются на притягивающий центр, а электроны — на нейтральный.

Время жизни дырок. Анализ уравнения

$$\frac{\partial}{\partial \delta} \frac{1}{\tau_p} = 0, \tag{13}$$

определяющего экстремальные точки функции $\tau_p(N)$, показывает, что при отчетливой немонотонности этой функции

$$\sqrt{2N_D} \gg \sqrt{n_t} \qquad \text{и} \qquad 3n_t^2 \ll 4\theta n_t^2 \ll \theta N_D^2, \tag{14}$$

где n_i — собственная концентрация свободных носителей. Для определения положения минимума представим уравнение (13) в форме

$$\delta^2 + 2\theta B\delta - \theta AB[1 + \Lambda_{1p}(\delta)] = 0, \tag{15}$$

где абсолютное значение функции

$$\Lambda_{1p}(\delta) = \frac{B + (A+\theta B)\theta}{\theta AB\delta} - \frac{B}{A\delta} +$$

$$+ \frac{B\delta^3 + (\delta^2 + B)[2B + (A+B)\delta]}{\theta AB\delta^6} \left[B + (A+\theta B)\delta + \theta B\delta^2 + \delta^3\right] +$$

$$+ \frac{B + A\delta - \delta^2}{\theta A\delta^5} \left\{A + (2\theta - 1)B + 4\theta B\delta + [3 + \theta(A+B)]\delta^2 - 2(\theta-1)\delta^3\right\} \tag{16}$$

много меньше единицы при

$$\delta = \delta_{1p}^{(0)} \equiv -\theta B + \sqrt{\theta B(A+\theta B)}. \tag{17}$$

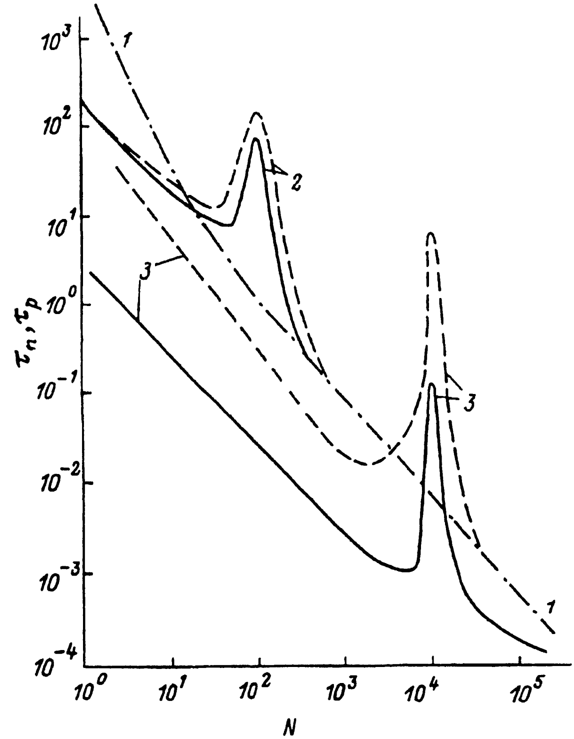

Рис. 1. Зависимость времен жизни дырок τ_p (сплошные линии) и электронов τ_n (штриховые) от концентрации рекомбинационных центров N при различных значениях концентрации мелких доноров N_D. N_D/n_i: *1* — <1, *2* — 10^2, *3* — 10^4. Времена жизни измерены в единицах $(n_i w_p)^{-1}$, концентрация N — в единицах n_i. Принято: $\theta = 10^2$, $n_t = 10^2 n_i$.

Это означает, что первый корень уравнения (13) δ_{1p} можно искать методом последовательных приближений, используя малый параметр $\Lambda_{1p}(\delta_{1p})$. Нулевое приближение (17) приводит к формуле

$$N_{1p} = N_D + 2\theta p_t - \sqrt{2\theta p_t(N_D + 2\theta p_t)} \tag{18}$$

для концентрации рекомбинационных центров $N = N_{1p}$, при которой зависимость $\tau_p(N)$ достигает минимума $\tau_{p\,\min}$ (рис. 1). Из этой формулы следует, что величина отношения N_{1p} к N_D увеличивается с увеличением N_D в интервале значений от 1/2 при $N_D \ll 2\theta p_t$ до $1 - \sqrt{2\theta p_t/N_D} \simeq 1$ при $N_D \gg 2\theta p_t$ (рис. 2). Далее показано, что экстремум типа максимума функции $\tau_p(N)$, как и функции $\tau_n(N)$, может иметь место только при значениях N, близких к N_D. Поэтому в выражении для верхнего предела значения N_{1p} сохранена малая поправка, которая, как показывает анализ, является основной.

Для определения положения максимума функции $\tau_p(N)$ преобразуем уравнение (13) к виду

$$[1 - \Lambda_{2p}(\delta)]\delta^2 = A + B, \tag{19}$$

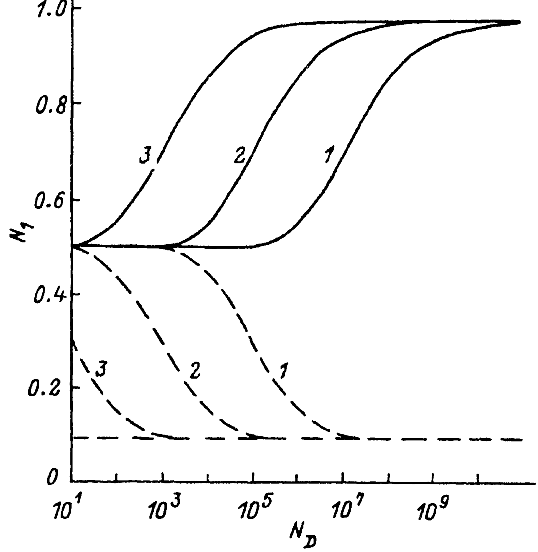

Рис. 2. Зависимость положений минимумов N_1 функций $\tau_p(N)$ (сплошные линии) и $\tau_n(N)$ (штриховые) от концентрации мелких доноров N_D при различных расположениях рекомбинационного уровня. n_t/n_i: *1* — 10^{-4}, *2* — 10^{-2}, *3* — 1, *4* — 10^2. Концентрация N_1 измерена в единицах N_D, концентрация N_D — в единицах n_i. Принято: $\theta = 10^2$.

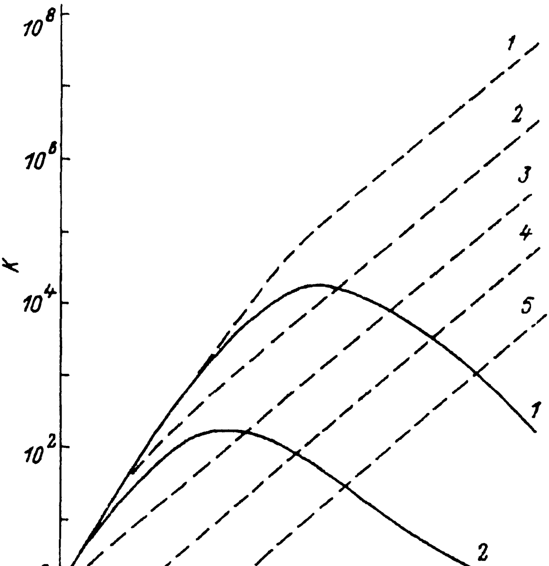

Рис. 3. Зависимость отношений K максимального к минимальному времени жизни дырок (сплошные линии) и электронов (штриховые) от концентрации мелких доноров N_D при различных расположениях рекомбинационного уровня. n_t/n_i: *1* — 10^{-4}, *2* — 10^{-2}, *3* — 1, *4* — 10^2, *5* — 10^4. Концентрация N_D измерена в единицах n_i. Принято: $\theta = 10^2$.

где абсолютное значение функции

$$\Lambda_{2p}(\delta) = \frac{1}{\theta\delta^4}\left[A + (2\theta - 1)B + 4\theta B\delta + 3\delta^2 - 2(\theta - 1)\delta^3\right] +$$

$$+ \frac{B + \delta^2}{\theta B\delta^5}\left[2B + (A + B)\delta + \delta^3\right]\left[1 + \frac{\delta(\delta + 1)(\theta B + \delta)}{B + A\delta - \delta^2}\right] \tag{20}$$

много меньше единицы при

$$\delta = \delta_{2p}^{(0)} \equiv \sqrt{A + B}. \tag{21}$$

Поэтому второй корень уравнения (13) δ_{2p} можно искать методом последовательных приближений, используя малый параметр $\Lambda_{2p}(\delta_{2p})$. Из соотношений (9) и (21) следует, что концентрация рекомбинационных центров $N = N_{2p}$, при которой зависимость $\tau_p(N)$ достигает максимума $\tau_{p\,\max}$ (рис. 1), в первом приближении по малому параметру $\Lambda_{2p}(\delta_{2p})$ определяется выражением

$$\frac{N_{2p}}{N_D} = 1 - \frac{1}{A + B} - \frac{\sqrt{A + B}}{A}\Lambda_{2p}(\sqrt{A + B}) \simeq 1. \tag{22}$$

В частности,

$$\frac{N_{2p}}{N_D} = \begin{cases} 1 - \dfrac{n_t}{2N_D}\left(1 + \dfrac{2}{\theta} + \dfrac{\sqrt{2N_D n_t}}{\theta p_t}\right) & \text{при} \quad N_D \gg 2p_t, \\ 1 - \left(\dfrac{2n_i}{N_D}\right)^2 & \text{при} \quad 2n_i \ll N_D \ll 2p_t. \end{cases} \tag{23}$$

Из соотношений (11), (12), (14), (17) и (21) следует, что

$$K_p(N_D) \equiv \frac{\tau_{p\,\max}}{\tau_{p\,\min}} = \begin{cases} N_D/8n_i & \text{при} \quad 8n_i \ll N_D \ll 2p_t, \\ \sqrt{N_D/32n_i} & \text{при} \quad N_D \ll 2\theta p_t \ll \theta N_D, \\ \sqrt{2}\theta p_t/\sqrt{N_D n_t} & \text{при} \quad \sqrt{2N_D n_t} \ll 2\theta p_t \ll N_D, \\ \sim 1 & \text{при} \quad 2\theta p_t \ll \sqrt{2N_D n_t}. \end{cases} \tag{24}$$

Отсюда видно, что функция $K_p(N_D)$ является немонотонной и может изменяться на несколько порядков (рис. 3). Значение

$$\tau_{p\,\max} = \left(1 + \frac{\sqrt{2}\theta p_t}{\sqrt{N_D n_t + 2n_i^2}}\right)(N_D w_p)^{-1} \tag{25}$$

увеличивается при уменьшении энергии рекомбинационного уровня $\mathcal{E}_t$ (рис. 4).

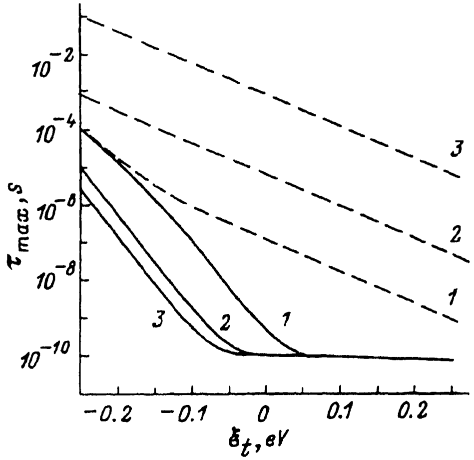

Рис. 4. Зависимость максимального времени жизни $\tau_{\max}$ дырок (сплошные линии) и электронов (штриховые) для Ge (*1*), Si (*2*) и GaAs (*3*) от положения рекомбинационного уровня по отношению к середине запрещенной зоны полупроводника $\Delta = \mathcal{E}_t - \mathcal{E}_g/2$ при $T = 300$ K. Использованы значения параметров полупроводников, приведенные в монографии [1]. Принято: $N_D = 10^{16}$ см$^{-3}$, $w_n = 10^{-8}$ см^3/с, $\theta = 10^2$.

Характер зависимости $\tau_{p\,\max}$ от температуры T определяется положением рекомбинационного уровня в запрещенной зоне полупроводника (рис. 5). При уменьшении температуры $\tau_{p\,\max}$ падает, если $\mathcal{E}_t \geqslant \mathcal{E}_g/2$, и растет, если $\mathcal{E}_t \leqslant \mathcal{E}_g/3$, где $\mathcal{E}_t$ отсчитывается от потолка валентной зоны, $\mathcal{E}_g$ — ширина запрещенной зоны. Если $\mathcal{E}_g/3 < \mathcal{E}_t < \mathcal{E}_g/2$, то зависимость $\tau_{p\,\max}(T)$ имеет максимальное значение при

$$T = \tilde{T} \equiv \frac{\mathcal{E}_t}{k}\,\ln^{-1}\left(2\,\frac{N_v}{N_D}\,\frac{\mathcal{E}_g - 2\mathcal{E}_t}{3\mathcal{E}_t - \mathcal{E}_g}\right), \tag{26}$$

где k — постоянная Больцмана, N_v — эффективная плотность состояний в валентной зоне.

Время жизни электронов. Анализ уравнения

$$\frac{\partial}{\partial\delta}\,\frac{1}{\tau_n} = 0, \tag{27}$$

определяющего экстремальные точки функции $\tau_n(N)$, показывает, что при отчетливой немонотонности этой функции

$$\sqrt{N_D} \gg \sqrt{2n_t} \qquad \text{и} \qquad 2n_t \gg N_D. \tag{28}$$

Для определения положения минимума представим уравнение (27) в виде

$$(\theta - 1)\delta^2 - 2\theta(A + B)\delta + \theta A(A + B)[1 + \Lambda_{1n}(\delta)] = 0, \tag{29}$$

где абсолютное значение функции

$$\Lambda_{1n}(\delta) = \frac{(\theta - 1)B}{\theta A(A + B)} - \frac{2B + A\delta}{\theta A(A + B)\delta}\left[\left(1 + \frac{2\theta B}{\delta}\right)\left(1 + \frac{A + B}{\delta^2} + \frac{2B}{\delta^3}\right) + \right.$$
$$\left. + \frac{1}{\delta}\left(A + B + \frac{2B}{\delta}\right)\left(\theta\frac{A + B}{\delta} - \theta + 1\right)\right] + \frac{B + A\delta - \delta^2}{\theta A(A + B)\delta}\times$$
$$\times\left[2 + (\theta - 1)\,\frac{A + B}{\delta} + \frac{6\theta B}{\delta} + \frac{2B}{\delta^2}\left(2\theta - 2 - \frac{1}{\delta}\right)\right] \tag{30}$$

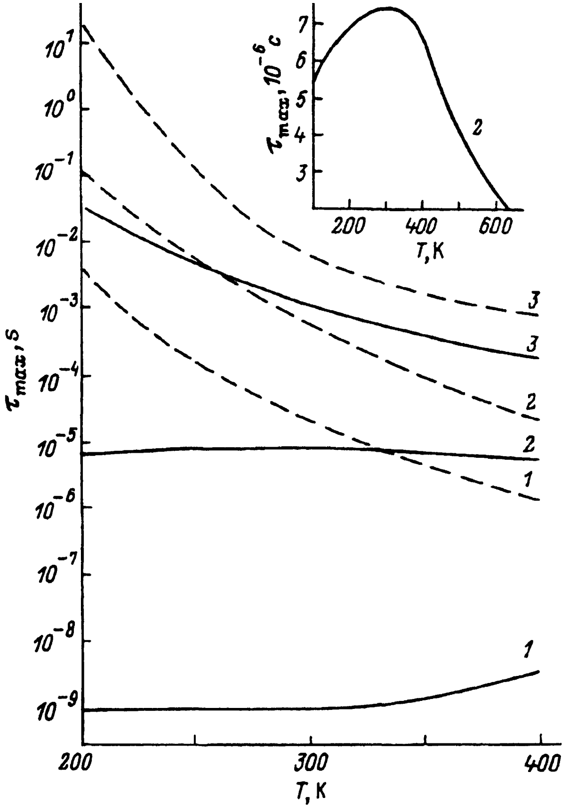

Рис. 5. Зависимость максимального времени жизни $\tau_{\max}$ дырок (сплошные линии) и электронов (штриховые) от температуры T для Si при различных положениях рекомбинационного уровня. Энергия рекомбинационного уровня $\mathcal{E}_t$ отсчитывается от потолка валентной зоны. Использованы значения параметров кремния, приведенные в монографии [1]. $\mathcal{E}_t$: *1* — $\mathcal{E}_g/2$, *2* — $(51/152)\mathcal{E}_g$, *3* — $\mathcal{E}_g/4$. Принято: $N_D = 10^{15}$ см$^{-3}$, $w_n = 10^{-8}$ см^3/с, $\theta = 10^2$.

много меньше единицы при

$$\delta = \delta_{1n}^{(0)} \equiv \frac{\sqrt{\theta(A+B)}}{\theta - 1}\left[\sqrt{\theta(A+B)} - \sqrt{A + \theta B}\right]. \tag{31}$$

Следовательно, первый корень уравнения (27) δ_{1n} можно искать, используя малый параметр $\Lambda_{1n}(\delta_{1n})$. Нулевое приближение (31) приводит к следующей формуле для концентрации рекомбинационных центров $N = N_{1n}$, при которой зависимость $\tau_n(N)$ достигает минимума $\tau_{n\,\min}$:

$$N_{1n} = \frac{\sqrt{N_D + 2\theta p_t}}{\theta - 1}\left[\sqrt{\theta(N_D + 2p_t)} - \sqrt{N_D + 2\theta p_t}\right]. \tag{32}$$

Из этой формулы следует, что величина отношения N_{1n} к N_D уменьшается с увеличением N_D в интервале значений от $1/2$ при $N_D \ll 2p_t$ до $1/(\sqrt{\theta}+1)$ при $N_D \gg 2\theta p_t$ (рис. 2).

Уравнение (27) можно преобразовать к виду (19), где вместо $\Lambda_{2p}(\delta)$ будет стоять функция

$$\Lambda_{2n}(\delta) = \frac{2\delta}{A} + \frac{2B}{A}\,\frac{\theta(A+B)\delta + 2\theta B + \delta}{\theta(A+B)\delta^4}\left(2A + B + \frac{2B}{\delta}\right) +$$

$$+ (2\theta B + \delta)\,\frac{A(A+B+\delta^2) + 2B\delta}{\theta A(A+B)\delta^3} - 2\,\frac{B + A\delta - \delta^2}{\theta A(A+B)\delta^4}\left(\delta^3 + 3\theta B\delta^2 - 2B\right) -$$

$$- (\theta - 1)\frac{2B[4B+(4A+B)\delta+(\delta-2)\delta^2]+2A(A+B)\delta^2+\delta(B-\delta^2)(A+B-\delta^2)}{\theta A(A+B)\delta^4}. \tag{33}$$

При значении $\delta = \delta_{2n}$, отвечающем максимуму $\tau_{n\,\max}$ функции $\tau_n(N)$, $|\Lambda_{2n}(\delta)| \ll 1$. Поэтому в нулевом приближении по $\Lambda_{2n}(\delta_{2n})$ значение $\delta_{2n} = \delta_{2p}^{(0)}$, концентрация $N = N_{2n}$, при которой $\tau_n(N) = \tau_{n\max}$, равна (как и в случае дырок) N_D, а

$$K_n(N_D) \equiv \frac{\tau_{n\max}}{\tau_{n\min}} = \begin{cases} N_D/8n_t & \text{при} \quad 2n_t \ll N_D \ll 2p_t, \\ (1/\varkappa)\sqrt{N_D/2n_t} & \text{при} \quad N_D \gg 2p_t, \end{cases} \tag{34}$$

где $\varkappa = 4$ при $\theta = 1$ и $\varkappa = 1$ при $\theta \gg 1$. Из (34) видно, что функция $K_n(N_D)$ в отличие от функции $K_p(N_D)$ монотонно растет с увеличением N_D и этот рост может составлять много порядков (рис. 3).

Значение

$$\tau_{n\max} = \sqrt{\frac{N_D + 2p_t}{2n_t}}\cdot(N_D w_n)^{-1} \tag{35}$$

увеличивается, как и $\tau_{p\max}$, при уменьшении энергии рекомбинационного уровня (рис. 4) и в отличие от $\tau_{p\max}$ всегда падает с увеличением температуры (рис. 5).

4. Физическая интерпретация

Дырки. Обратное время жизни дырок,

$$\frac{1}{\tau_p} = \sum_{i=1}^{3}\frac{1}{\tau_{pi}}, \tag{36}$$

складывается из трех парциальных составляющих. Первая их них,

$$\frac{1}{\tau_{p1}} = w_p N_-^e = w_p N_D\,\frac{B + A\delta - \delta^2}{A\delta}, \tag{37}$$

отвечает изменению скорости захвата дырок $\Delta p w_p N_-^e$, обусловленному лишь отклонением концентрации дырок от своего равновесного значения (захват избыточных дырок Δp на равновесные центры захвата N_-^e). Вторая составляющая,

$$\frac{1}{\tau_{p2}} = -w_p N_D\,\frac{B}{A\delta}\,\frac{(\delta\theta - 1)(B + A\delta - \delta^2)}{A\delta + B[1 + \delta\theta(1+\delta)] + \delta^3}, \tag{38}$$

отвечает изменению скорости захвата дырок $p_e w_p \Delta N_-$, обусловленному лишь отклонением концентрации центров захвата дырок от своего равновесного значения (захват равновесных дырок p_e на неравновесные центры захвата ΔN_-). Третья составляющая,

$$\frac{1}{\tau_{p3}} = \frac{\delta}{\tau_{p2}}, \tag{39}$$

отвечает изменению скорости теплового выброса дырок с примесного уровня в валентную зону $2p_t w_p \Delta N_0$, обусловленному отклонением концентрации центров генерации дырок от своего равновесного значения (тепловая генерация дырок с неравновесных центров ΔN_0). Время τ_{p1} можно рассматривать как время захвата избыточных дырок равновесными ловушками, время τ_{p2} — как время релаксации избыточных дырок за счет захвата равновесных дырок неравновесными ловушками, а время τ_{p3} можно считать временем термической генерации дырок с неравновесных центров.

При выполнении условий (14) и значениях N, меньших N_D, рекомбинационные центры практически полностью заполнены электронами ($\delta \gg 1$). По этой причине даже при $\theta = 1$ время захвата дырки τ_{p1} много меньше времени захвата электрона τ_{n1} на соответствующие равновесные центры захвата. Другими словами, дырки захватываются равновесными центрами гораздо более интенсивно, чем электроны. Следовательно, появление избыточных свободных носителей приво дит к появлению дополнительных, неравновесных центров тепловой генерации дырок и, одновременно, к уменьшению их центров захвата ($\Delta N_0 = -\Delta N_- > 0$). При таком изменении зарядового состояния атомов рекомбинационной примеси составляющие $1/\tau_{p2}$ и $1/\tau_{p3}$ в выражении (36) имеют отрицательные значения, причем $|\tau_{p3}| \ll |\tau_{p2}|$, так как при $N < N_D$ $\delta \gg 1$. Это означает, что в рассматриваемых условиях преимущественно за счет тепловой генерации дырок с их неравновесных центров генерации время жизни дырок τ_p превосходит время их захвата на равновесные ловушки τ_{p1}. Пока $N < N_D$, число равновесных центров захвата дырок растет ($N_-^e \simeq N$) при увеличении значения N, а центров захвата электронов остается по-прежнему мало ($N_0^e \ll N$). Поэтому растет и концентрация неравновесных центров тепловой генерации дырок ΔN_0. Этот рост вызывает более быстрое уменьшение $|\tau_{p3}|$, чем уменьшение τ_{p1}. В результате, начиная с концентрации $N = N_{1p}$, скорость тепловой генерации дырок с неравновесных центров и скорость захвата неравновесных дырок на равновесные ловушки становятся близки друг к другу. По этой причине τ_p начинает расти (рис. 1).

Когда N становится больше чем N_D, равновесная концентрация дырочных ловушек N_-^e практически стабилизируется, в то время как равновесная концентрация электронных ловушек N_0^e возрастает с ростом концентрации рекомбинационных центров ($N_0^e \simeq N - N_D$ при $n_t/2 \ll N - N_D \ll N_D^2/2p_t$, $N_-^e \simeq N_D$). Это означает, что отношение τ_{p1}/τ_{n1} увеличивается с увеличением N. По этой причине концентрация неравновесных центров термической генерации дырок уменьшается, а концентрация дырочных ловушек N_- возрастает. В результате $\tau_p(N)$ уменьшается с увеличением N (рис. 1). Когда N становится больше чем $N_D^2/2p_t$, равновесная концентрация дырочных ловушек снова возрастает с ростом N за счет термического выброса электронов из валентной зоны на примесный уровень ($N_-^e \simeq \sqrt{2p_+N}$). Однако равновесная концентрация электронных ловушек возрастает гораздо быстрее ($N_0^e \simeq N$). Поэтому уменьшение $\tau_p(N)$ продолжается. Как видно из выражений (36)–(39), τ_p становится меньше чем τ_{p1}, когда $\delta\theta$ становится меньше единицы.

Как было показано выше, с увеличением N_D положение минимума N_{1p} функции $\tau_p(N)$ сдвигается к N_D (рис. 2). Причина заключается в том, что равновесные электроны, захватываясь неравновесными центрами тепловой генерации дырок, уменьшают концентрацию этих центров ΔN_0. Чем больше равновесная концентрация электронов n_e, тем больше это уменьшение. Значение n_e растет с ростом N_D. Когда N возрастает, n_e уменьшается, а N_-^e увеличивается, что вызывает увеличение ΔN_0 при $N < N_D$. Другими словами, уменьшение ΔN_0 при увеличении N_D компенсируется увеличением ΔN_0 при увеличении N. Вот почему, чем больше N_D, тем ближе N_{1p} к N_D.

По тем же самым причинам немонотонная зависимость τ_p от N исчезает, как было показано выше, при $N_D > 2(\theta p_t)^2/n_t$ (рис. 3). Немонотонная зависимость τ_p от N не имеет места и при малых значениях N_D (см. неравенства (14), рис. 1,3), когда степень равновесной заселенности электронами рекомбинационного уровня определяется в основном электронно-дырочными переходами между ним и свободными зонами. В этом случае значения величины δ не могут обеспечить более быстрый рост скорости тепловой генерации дырок с неравновесных центров по сравнению со скоростью захвата неравновесных дырок на их равновесные ловушки при увеличении N. Максимальное значение $\tau_{p\,\max}/\tau_{p\,\min}$ примерно равно $\sqrt{\theta p_t/n_t}$ и достигается при $N_D \simeq 2\theta p_t$.

Отметим также, что по мере увеличения энергии рекомбинационного уровня $\mathcal{E}_t$ немонотонная зависимость $\tau_p(N)$ постепенно ослабевает (рис. 3) и затем исчезает совсем. Это обусловлено увеличением концентрации равновесных электронов, уменьшением значения δ и вероятности теплового выброса дырки с рекомбинационного уровня в валентную зону по мере удаления его положения от потолка валентной зоны. По той же причине значение $\tau_{p\,\max}$ уменьшается при увеличении $\mathcal{E}_t$ (рис. 4).

Характер зависимости $\tau_{p\,\max}$ от температуры (рис. 5) определяется температурными зависимостями

$$\delta_{2p}(T) = \frac{1}{n_i(T)} \sqrt{2N_D p_t(T) + 4p_t^2(T)}, \tag{40}$$

$$p_t(T) = N_v \exp\left(-\frac{\mathcal{E}_t}{kT}\right) \quad \text{и} \quad n_e(T) = \frac{\delta_{2p}(T) n_i^2(T)}{2p_t(T)}. \tag{41}$$

Значения $p_t(T)$ и $n_e(T)$ всегда растут с ростом T. Увеличение p_t означает увеличение вероятности теплового выброса дырки с рекомбинационного уровня в валентную зону. Поэтому упомянутый процесс способствует росту $\tau_{p\,\max}$ при увеличении T. Рост $n_e(T)$, вместе с тем, способствует падению $\tau_{p\,\max}$ при увеличении T за счет уменьшения концентрации неравновесных центров тепловой генерации дырок ΔN_0.

Значение δ_{2p} падает с ростом T при $\mathcal{E}_t \leqslant \mathcal{E}_g/2$ за счет все более близкого приближения N_{2p} к N_D (см. выражение (23)). Оно падает и при $\mathcal{E}_t > \mathcal{E}_g/2$ вплоть до температур, при которых немонотонные зависимости τ_p и τ_n от N исчезают. Падение значения δ_{2p} уменьшает ΔN_0, а поэтому способствует уменьшению $\tau_{p\max}$ с ростом температуры. Когда $\mathcal{E}_t \geqslant \mathcal{E}_g/2$, то $p_t(T)$ растет быстрее, чем падает $\delta_{2p}(T)$ и растет $n_e(T)$. В результате $\tau_{p\max}$ увеличивается с увеличением температуры. Если $\mathcal{E}_t \leqslant \mathcal{E}_g/3$, то рост p_t с увеличением T не может компенсировать уменьшение δ_{2p} и увеличение n_e. В результате $\tau_{p\max}$ падает с увеличением температуры. Если $\mathcal{E}_g/3 < \mathcal{E}_t < \mathcal{E}_g/2$, то по тем же самым причинам, что и в предыдущих случаях, при $T < \tilde{T}$ функция $\tau_{p\max}(T)$ — растущая, а при $T > \tilde{T}$ — падающая.

Электроны. Аналогично ситуации с дырками, обратное время жизни электронов складывается из трех парциальных составляющих:

$$\frac{1}{\tau_{n1}} = w_n N_0^e = w_n N_D \frac{B + A\delta - \delta^2}{A\delta^2}, \tag{42}$$

$$\frac{1}{\tau_{n2}} = w_n N_D \frac{\delta\theta - 1}{A} \frac{B + A\delta - \delta^2}{\theta(B + A\delta - \delta^2) + (1+\delta)(\theta B + \delta)}, \tag{43}$$

$$\frac{1}{\tau_{n3}} = \frac{1}{\delta\tau_{n2}}. \tag{44}$$

Времена τ_{n1}, τ_{n2} и $-\tau_{n3}$ имеют физический смысл, подобный временам τ_{p1}, τ_{p2} и $-\tau_{p3}$ соответственно.

Когда $N < N_D$, величина $\delta \gg 1$, и, следовательно, $\tau_{p1} \ll \tau_{n1}$. Поэтому появление избыточных свободных носителей приводит к возникновению дополнительных, неравновесных центров захвата электронов и одновременно к уменьшению центров их генерации ($\Delta N_0 = -\Delta N_- > 0$). При таком изменении зарядового состояния рекомбинационных центров парциальные составляющие $1/\tau_{n2}$ и $1/\tau_{n3}$ имеют положительные значения, причем $\tau_{n2} \ll \tau_{n3}$, так как при $N < N_D$ величина $\delta \gg 1$. Это означает, что в рассматриваемых условиях, преимущественно за счет захвата равновесных электронов на неравновесные ловушки, время жизни электронов τ_n меньше времени их захвата на равновесные ловушки τ_{n1}. При увеличении значения N число равновесных центров захвата дырок растет, а центров захвата электронов остается по-прежнему малым. В результате растет концентрация неравновесных центров захвата электронов ΔN_0. По этой причине, начиная с концентрации $N = N_{1n}$, скорость захвата равновесных электронов на неравновесные центры становится больше скорости захвата неравновесных электронов на равновесные ловушки. Другими словами, определяющей составляющей обратного времени жизни электронов $1/\tau_n$ становится $1/\tau_{n2}$, которая падает с ростом N за счет уменьшения концентрации равновесных электронов n_e, т.е. $\tau_n \simeq \tau_{n2}$, и растет с увеличением N (рис. 1).

При значениях N, бóльших N_D, отношение τ_{p1} к τ_{n1} растет с увеличением N. Это в свою очередь приводит к уменьшению концентрации неравновесных центров захвата электронов при увеличении N. Кроме того, продолжает падать значение n_e. В результате определяющей составляющей становится $1/\tau_{n1}$, и поэтому $\tau_n(N)$ падает с ростом N (рис. 1).

Как было показано выше, в отличие от зависимости $\tau_p(N)$ отношение N_{1n} к N_D уменьшается (рис. 2), а $\tau_{n\,\max}/\tau_{n\,\min}$ всегда монотонно возрастает с ростом N_D (рис. 3). Эти закономерности обусловлены ростом n_e при увеличении N_D. За счет этого скорость захвата равновесных электронов неравновесными ловушками становится больше скорости захвата неравновесных электронов равновесными ловушками при меньших концентрациях ΔN_0, т. е. при меньших значениях N/N_D. В противоположность ситуации с дырками, уменьшение ΔN_0 компенсируется увеличением n_e.

Подобно дырочному времени жизни по мере уменьшения N_D или увеличения $\mathcal{E}_t$ немонотонная зависимость τ_n от N постепенно ослабевает и затем исчезает совсем (рис. 1, 3). Первая из этих закономерностей связана с тем, что с уменьшением N_D уменьшаются n_e и значение величины $\delta = 2n_e/n_t$. Вторая закономерность обусловлена уменьшением значения δ, а следовательно, и ΔN_0 с увеличением энергии рекомбинационного уровня $\mathcal{E}_t$. При этом, однако, за счет роста n_e немонотонность исчезает при бóльших значениях $\mathcal{E}_t$, чем в случае дырок.

За счет уменьшения δ_{2n} уменьшается и значение $\tau_{n\,\max}$ при увеличении $\mathcal{E}_t$ (рис. 4). Характер зависимости $\tau_{n\,\max}$ от температуры (рис. 5) определяется только зависимостью $\delta_{2n}(T)$, поскольку в точке максимума $\tau_n = \tau_{n1}/2 \sim 1/N_0^e \sim \delta_{2n}$. В нулевом приближении $\delta_{2n}(T)$ совпадает с $\delta_{2p}(T)$, определяемым выражением (40). Поэтому $\tau_{n\,\max}$ всегда уменьшается с увеличением температуры.

5. Заключение

При увеличении концентрации рекомбинационных центров растет темп захвата ими неравновесных носителей. Тем не менее, как показано в данной работе, за счет слабо неравновесного заполнения рекомбинационного уровня это совсем не обязательно должно сопровождаться уменьшением времен жизни неравновесных носителей. Дело в том, что времена жизни определяются не только захватом неравновесных носителей на равновесные ловушки, но и тепловым выбросом электронов и дырок неравновесными центрами из связанного состояния в свободное, а также захватом равновесных носителей неравновесными ловушками. По этой причине времена жизни неравновесных носителей могут оказаться как больше, так и меньше времени их захвата на равновесные ловушки и быть сильно немонотонными функциями концентрации центров рекомбинации. Существенно, что отношение времен жизни в максимуме и минимуме этих функций может составлять несколько порядков. Рассмотренные в данной работе закономерности реализуются при сколь угодно низких уровнях возбуждения и проявляются тем ярче, чем больше ширина запрещенной зоны полупроводника.

В статье [11] сообщалось о наличии минимума и участка слабого роста (на 24%) в экспериментальной зависимости времени жизни неравновесных носителей от концентрации рекомбинационных центров, которая увеличивалась вследствие облучения образца электронами высоких энергий.

Результаты данной работы также показывают, что фоточувствительность полупроводников можно сильно увеличить и при большой концентрации рекомбинационных центров, если оптимально выбрать концентрацию мелкой легирующей примеси и энергетическое положение рекомбинационного уровня в запрещенной зоне полупроводника.

Тематика данной работы поддержана Российским фондом фундаментальных исследований.

Список литературы

[1] С. Зи. *Физика полупроводниковых приборов* (М., Мир, 1984) ч. 1, 2.
[2] А. Милнс. *Примеси с глубокими уровнями в полупроводниках* (М., Мир, 1982).
[3] R.N. Hall. Phys. Rev., **87**, 387 (1952).
[4] W. Shockley, W.T. Read. Phys. Rev., **87**, 835 (1952).
[5] Дж.С. Блекмор. *Статистика электронов в полупроводниках* (М., Мир, 1964).
[6] С.М. Рывкин. *Фотоэлектрические явления в полупроводниках* (М., Физматгиз, 1963).
[7] P.M. Downey, R.J. Martin, K.E. Nahory, O.G. Lorimor. Appl. Phys. Lett., **46**, 396 (1985).
[8] A. Schaelin, R. Loepfe, H. Melchior, M. Suter, W. Woelfi. Mater. Sci. Engin. B, **2**, 81 (1989).
[9] А.А. Другова, В.А. Холоднов. Письма ЖТФ, **18**, 23 (1992).
[10] В.В. Осипов, В.А. Холоднов. ФТП, **4**, 2241 (1970).
[11] G.K. Wertheim. Phys. Rev., **109**, 1086 (1958).
[12] Р. Смит. *Полупроводники* (М., Мир, 1982).
[13] Дж. Блейкмор. *Физика твердого тела* (М., Мир, 1988).
[14] В.А. Холоднов. Автореф. канд. дис. (М., МФТИ, 1973).
[15] П.С. Киреев. *Физика полупроводников* (М., Высш. шк., 1969).

Редактор Л.В. Шаронова

To the theory of Hall–Shockley–Reed recombination

V.A. Kholodnov

State Research Center of Russia, Research and Production Enterprize «Orion», 111123 Moscow, Russia

A study has been made of how the value of the concentration of recombination centers N affects the Hall–Shockley–Reed generation–recombination process. A small departure of the semiconductor from equilibrium has been considered. It is shown that the concentration dependences of non–equilibrium electron and hole lifetimes, $\tau_n(N)$ and $\tau_p(N)$, respectively, which, on the whole, diminish with the N increase, under certain conditions exhibit a portion of growth rate by several orders of magnitude greater. Analytical solution to the extremum problem of $\tau_n(N)$ and $\tau_p(N)$ functions is found. A careful consideration is given to the dependence of the localization of the extrema as well as the function values in the extremum points on the shallow doping impurity density N_D, the energy of the recombination level $\mathcal{E}_t$, the ratio of the hole and electron capture probabilities θ, the bandgap energy $\mathcal{E}_g$ and temperature T. A detailed physical interpretation of results obtained is suggested.