

Генерационно-рекомбинационные процессы в полупроводниках

© И.Н. Воловичев*[†], Ю.Г. Гуревич[†]

* Институт радиофизики и электроники Национальной академии наук Украины, 310085 Харьков, Украина

[†] Departamento de Física, CINVESTAV-IPN, 07000 México, D.F., México

(Получена 22 июня 2000 г. Принята к печати 16 августа 2000 г.)

Сформулирован единый методический подход к изучению явлений переноса в полупроводниках. Проанализированы различные модели рекомбинационных процессов, используемые при изучении транспортных явлений и установления равновесия в полупроводниковых структурах. Приведены новые выражения, описывающие рекомбинационные процессы в стационарных режимах в произвольных температурных полях. Проанализирован процесс рекомбинации в теории горячих носителей, когда температуры носителей тока и фононов не совпадают между собой. Изучены проявления условия квазинейтральности в термодинамическом равновесии и в явлениях переноса.

1. Введение

В связи с широким интересом к изучению транспортных явлений в полупроводниках и полупроводниковых структурах исследователям приходится часто сталкиваться с таким явлением, как неравновесные носители тока. Их появление может быть обусловлено различными факторами — инжекцией основных или неосновных носителей, освещением с энергией фотонов большей ширины запрещенной зоны (фотоэффект) и др. В последнее время было показано, что появление неравновесных носителей, влияющих на токоперенос, возможно уже в линейном приближении по температурному и электрическому полям. Оно связано с перераспределением носителей в этих полях [1]. Во всех этих задачах принципиальную роль играет рекомбинация неравновесных носителей в образце. Концентрации последних определяются хорошо известными уравнениями непрерывности:

$$\frac{\partial n}{\partial t} = g_n + \frac{1}{e} \operatorname{div} \mathbf{j}_n - R_n, \quad \frac{\partial p}{\partial t} = g_p - \frac{1}{e} \operatorname{div} \mathbf{j}_p - R_p, \quad (1)$$

где $n = n_0 + \delta n$ и $p = p_0 + \delta p$ — концентрация электронов и дырок, δn и δp — концентрация неравновесных носителей, $\mathbf{j}_{n,p}$ — парциальные токи электронов и дырок, $g_{n,p}$ — темп внешней генерации и $R_{n,p}$ — темп рекомбинации носителей. В свою очередь темп рекомбинации $R_{n,p} = r_{n,p} - g_{T_n, T_p}$ определяется разностью темпов двух конкурирующих процессов: захватом носителей $r_{n,p}$ и их тепловой генерацией g_{T_n, T_p} .

Ясно, что данная система уравнений в общем случае является неполной и должна быть дополнена уравнением Пуассона:

$$\operatorname{div} \mathbf{E} = \frac{4\pi}{\epsilon} \rho, \quad (2)$$

где \mathbf{E} — напряженность электрического поля в полупроводнике, ρ — плотность объемного заряда, ϵ — диэлектрическая проницаемость полупроводника.

Для применения системы уравнений (1)–(2) необходимо конкретизировать зависимости парциальных токов, плотности объемного заряда и темпов рекомбинации от

концентрации неравновесных носителей. К сожалению, в огромном числе работ, посвященных этим вопросам, имеют место некорректности, а зачастую прямые ошибки, связанные с тем, что сама рекомбинация (члены $R_{n,p}$ в уравнениях (1)) зачастую описывается некорректными выражениями, а именно

$$R_n = \delta n / \tau_n, \quad R_p = \delta p / \tau_p,$$

где $\tau_{n,p}$ — времена жизни неравновесных носителей, являющиеся параметрами рассматриваемого полупроводника [2–7]. Но, поскольку в статическом режиме должно выполняться условие непрерывности полного тока $\operatorname{div} \mathbf{j} = 0$, возникает дополнительное условие $\delta n / \tau_n = \delta p / \tau_p$, которое физически ниоткуда не следует и непонятно что описывает. Ряд авторов (см., например, монографию [2]) используют это условие для уменьшения числа переменных, что в корне неверно. Встречается также подход, когда данное условие рассматривается как условие, связывающее времена жизни носителей [3]. Данный метод, помимо нефизичности, является малопродуктивным, так как времена жизни носителей являются уже не параметрами полупроводника, а функциями концентрации неравновесных носителей, в свою очередь подлежащими определению, что влечет к появлению существенных математических трудностей.

Весьма распространен иной подход [8], при котором полагают $R_n = R_p = \delta p / \tau_p$, где δp и τ_p — соответственно концентрация неравновесных неосновных носителей и их время жизни. В этом случае условие непрерывности полного тока выполняется тождественно, однако имеет место иное серьезное противоречие. Оно становится особенно явным, если рассмотреть случай, когда имеет место инжекция основных носителей. Из физических соображений очевидно, что инжестрированные неравновесные основные носители рекомбинируют. В то же время, с формальной точки зрения, поскольку неравновесные носители не появляются ($\delta p = 0$), темп рекомбинации равен нулю.

Проблема резко усложняется, если температура в образце является неоднородной, т.е. зависит от координаты. При этом становится непонятным, от какого значения отсчитывается концентрация неравновесных носителей, т.е. что понимать под величинами n_0 и p_0 . Выход из положения был предложен в работе [9], однако данный метод является довольно искусственным и применим только в случае известного (фиксированного) температурного поля.

Еще одна проблема связана с уравнением Пуассона (2). Уже при вычислении встроенного электрического поля, появляющегося в равновесии в неоднородных материалах и полупроводниковых структурах, зачастую объемный заряд описывают выражениями, пригодными для анализа лишь отдельных частных случаев, не конкретизируя границы применимости используемого приближения. В целом же этот вопрос остается неисследованным. Такое положение тем более распространено, когда речь идет о неравновесной ситуации.

Все указанные ошибки и неточности, как нам представляется, возникают из-за того, что ни в одной известной нам монографии (или обзоре) не описан с общих позиций этот круг явлений. Иногда выписываются верхние исходные выражения, а потом делаются ничем не обоснованные и неконтролируемые приближения. Примером такого рода может служить широко распространенное утверждение о равенстве неравновесной концентрации электронов и дырок в каждой точке образца при межзонной рекомбинации [2].

Нам представляется целесообразным в рамках простейших моделей последовательно описать процедуру нахождения неоднородных равновесных и неравновесных концентраций носителей. Подчеркнем сразу, что большинство выводов, изложенных далее (но не все), можно найти в монографиях, обзорах и статьях. Однако, как правило, корректный учет одного фактора в указанных работах сопровождается пренебрежением других факторов, оказывающих на изучаемый эффект влияние такого же порядка.

2. Термодинамическое равновесие

Построение полностью самосогласованной общей схемы решения задач переноса тока удобно начать с рассмотрения простейшей ситуации — состояния термодинамического равновесия. Здесь и далее во избежание математических трудностей мы ограничимся линейным анализом.

Итак, пусть имеется произвольный однородный полупроводник, в котором концентрации электронов и дырок равны соответственно n_0 и p_0 . Поскольку в общем случае $n_0 \neq p_0$, в полупроводнике имеется также связанный заряд электронов с концентрацией n_i^0 на ионизованных примесных центрах. Суммарная концентрация примесных центров равна N_i . Образец нейтрален в каждой точке ($n_0 + n_i^0 - p_0 = 0$), следовательно, отсутствует

встроенное поле. Температура образца также однородна: $T_0 = \text{const}$. Уровни химических потенциалов электронов μ_n^0 и дырок μ_p^0 постоянны в пространстве ($\mu_{n,p}^0 = \text{const}$) и совпадают друг с другом ($\mu_n^0 = -\varepsilon_g - \mu_p^0$, где ε_g — ширина запрещенной зоны полупроводника), т.е., как и должно следовать из условия термодинамического равновесия, уровень электрохимического потенциала системы постоянен в пространстве и един для всех подсистем носителей.

Ситуацию в неоднородном полупроводнике можно описать исходя из изложенного выше исходного состояния однородного полупроводника. Для определенности, предположим, что неоднородность создана за счет неоднородного легирования:

$$N_i(x) = N_i^0 + \delta N_i(x).$$

Мысленно разобьем образец на тонкие слои, изолированные друг от друга (так, что перетекание носителей невозможно). После ионизации введенных примесей в каждом таком слое установятся "равновесные" концентрации носителей, которые имели бы место в массивном однородном полупроводнике с тем же значением концентрации примеси во всем объеме образца. Эти концентрации обозначим

$$n^0(x) = n_0 + \delta n(x), \quad p^0(x) = p_0 + \delta p(x)$$

и соответственно химические потенциалы

$$\mu_{n,p}^0(x) = \mu_{n,p}^0 + \delta \mu_{n,p}(x).$$

Поскольку в каждом слое электроны и дырки находятся в равновесии друг с другом, их уровни химических потенциалов совпадают:

$$\mu_n^0(x) = -\varepsilon_g - \mu_p^0(x).$$

Однако химические потенциалы уже неоднородны в пространстве, т.е. $\nabla \mu_{n,p} \neq 0$ — термодинамическое равновесие между слоями отсутствует. Тем не менее, поскольку перетекания носителей еще не произошло, объемный заряд в полупроводнике отсутствует.

Если теперь "соединить" рассматриваемые слои, то произойдет перетекание носителей, сопровождаемое возникновением встроенного электрического поля и выравниванием электрохимического потенциала электронов и дырок. Состояние термодинамического равновесия рассматриваемой неоднородной системы будет характеризоваться новым распределением носителей:

$$n(x) = n^0(x) + \delta n_1(x), \quad p(x) = p^0(x) + \delta p_1(x)$$

и в полупроводнике возникнет встроенное электрическое поле с потенциалом $\varphi(x)$. Уровень химического потенциала электронов и дырок изменится

$$\mu_n = \mu_n^0(x) + \delta \mu_n'(x), \quad \mu_p = \mu_p^0(x) + \delta \mu_p'(x),$$

но по-прежнему будет удовлетворять соотношению

$$\mu_n(x) = -\varepsilon_g - \mu_p(x).$$

Этот уровень не будет постоянен в пространстве, в то время как градиент электрохимического потенциала равен нулю $\nabla \tilde{\mu}_{n,p} = 0$ ($\tilde{\mu}_{n,p} = \mu_{n,p} \mp e\varphi(x)$).

Отметим, что на самом деле переходный процесс установления равновесия, вообще говоря, отличается от изложенной выше схемы, так как время установления равновесия в каждом слое (которое порядка времени жизни неравновесных носителей в данном полупроводнике) значительно превышает максвелловское время, за которое происходит перетекание носителей. Таким образом, перетекание носителей адиабатически отслеживает их генерацию (которая в данном случае является источником неравновесных носителей), т.е. эти процессы протекают одновременно и физически выделить две стадии переходного процесса невозможно. В случае же, когда неоднородность создается приведением достаточно тонкого полупроводника в контакт с другим материалом, первичным становится перетекание носителей под действием разности термодинамических работ выхода контактирующих веществ. Процесс формирования единого уровня химического потенциала начинается лишь на следующем этапе. Тем не менее изложенная (искусственная, мысленная) процедура разбиения переходного процесса на несколько стадий, подразумевающих наличие равновесия на каждой из них по отношению к какому-либо одному фактору, является довольно продуктивной, позволяя решать некоторые проблемы анализа транспортных процессов [9,10] или лучше понять некоторые особенности формирования равновесного состояния (например, изменение полного числа носителей в структуре с гетеропереходом [11]).

С физической точки зрения единство (но не постоянство, в общем случае) химического потенциала носителей ($\mu_n = -\varepsilon_g - \mu_p$) означает равенство скоростей процесса захвата носителей примесными центрами (или межзонной рекомбинации) и их термической генерации, т.е. равенство нулю скорости рекомбинации в образце, которая зависит от концентрации всех сортов частиц, участвующих в акте рекомбинации.

Таким образом, в состоянии равновесия

$$R_n(n, p, n_t) = 0, \quad (3)$$

$$R_p(n, p, n_t) = 0. \quad (4)$$

Третьим уравнением является условие баланса частиц на примесных уровнях. В статическом случае данное условие имеет следующий вид:

$$R_n(n, p, n_t) = R_p(n, p, n_t). \quad (5)$$

Обратим внимание, что уравнение (5), сформулированное на основе рассмотрения баланса частиц на

примесном уровне, одновременно обеспечивает тождественное выполнение вытекающего из уравнений Максвелла требования к любому статическому токопереносу $\text{div } \mathbf{j} = 0$.

Как легко заметить, в состоянии термодинамического равновесия уравнения (3)–(5) не являются линейно независимыми, однако вместо одного из них имеем условие равенства нулю парциальных токов носителей: $\mathbf{j}_n = \mathbf{j}_p = 0$. Отметим, что последнее есть ни что иное, как переформулированное в терминах парциальных токов условие постоянства электрохимического потенциала.

Четвертым уравнением, замыкающим рассматриваемую систему уравнений с четырьмя неизвестными, является уравнение Пуассона (2). Именно приведенная система уравнений является наиболее общей математической формулировкой задачи о термодинамическом равновесии неоднородного полупроводника.

Из нее вполне очевидным является переход к случаю однородного полупроводника. Действительно, если образец однороден, то электрическое поле и объемный заряд отсутствуют ($\mathbf{E} = 0$, $\rho = 0$), уравнение Пуассона выполняется тождественно и система уравнений состоит только из трех уравнений с тремя неизвестными n , p , n_t , т.е. данный предельный переход по-прежнему приводит к корректной математической постановке задачи.

Если же и концентрация зарядов на примесных уровнях (n_t) известна (задана), то происходит дальнейшее сокращение числа независимых переменных и уравнений до двух. Самосогласованность такого перехода обусловлена тем, что при корректно заданной концентрации зарядов на примесных уровнях условие $R_n(n, p, n_t) = R_p(n, p, n_t)$ выполняется тождественно по крайней мере для равновесных (искомых) концентраций носителей и переопределения системы уравнений не происходит.

Наконец, исследуем описываемый подход в часто используемом приближении квазинейтральности [12]. Предположим, что характерные параметры задачи (а в данном случае это размер образца) намного превышают радиус экранирования Дебая r_D основных носителей. В этом случае можно положить $\delta p = 0$ и уравнение Пуассона становится излишним. Если же, как указывалось выше, использовать широко применяемое предположение о неизменности концентрации зарядов на примесных уровнях (далее в разд. 4 будут приведены соответствующие критерии), то условие квазинейтральности сводится к равенству $\delta n = \delta p$. С другой стороны, единство уровня химического потенциала (т.е. отсутствие рекомбинации) требует $\delta \mu_n = -\delta \mu_p$ или для невырожденного полупроводника $\delta p = -(p_0/n_0)\delta n$. Последнее означает, что в монополярном полупроводнике (например, в электронном, где $n_0 \gg p_0$) объемным зарядом неосновных носителей можно пренебречь. Но в силу условия квазилинейности это означает, что и основные носители не создают объемный заряд: $\delta n = 0$. Аналогично для собственного полупроводника совместность

условий квазинейтральности и единства уровня химического потенциала возможно только при $\delta n = \delta p = 0$. Другими словами, для любого полупроводника квазинейтральность в равновесии означает неизменность концентрации носителей при любом воздействии извне (например, при приложении поляризующего поля или создании твердотельной структуры). Этот вывод хорошо понятен физически. Действительно, любое такое воздействие, оставляющее систему в состоянии равновесия (т.е. не вызывающее появления тока) и не изменяющее концентрацию заряженных примесей, будут экранироваться на длине порядка радиуса Дебая, т.е. в рассматриваемом приближении ($L \gg r_D$) — поверхностными зарядами.

3. Горячие носители в однородном образце

В качестве более сложной ситуации рассмотрим однородный биполярный полупроводник, в котором имеет место однородный разогрев одной из подсистем носителей (для определенности, электронов). Такой случай может иметь место при освещении полупроводника слабопоглощающимся светом с энергией фотонов, меньшей ширины запрещенной зоны, когда основным механизмом поглощения света является его поглощение свободными носителями.

В данном примере, как и в случае термодинамического равновесия, отсутствует как полный ток (цепь разомкнута), так и парциальные токи носителей. Однако имеет место энергетическая неравновесность: из-за разогрева электронов их температура (T_e) отличается от температуры дырок (T_p) и фононов (которые мы принимаем одинаковыми) $T_e \neq T_p = T_0$.

Рассмотрим процесс установления стационарного режима последовательно (в две стадии).

На первом этапе, в результате разогрева электронного газа, его концентрация остается неизменной, но уровень химического потенциала изменяется (ввиду явной его зависимости от температуры), оставаясь однородным в силу однородности разогрева. В линейном приближении это изменение равно

$$\delta\mu_n = \mu_n^0(T_e - T_0)/T_0 \quad (\mu_n'(T_e) = \mu_n^0 + \delta\mu_n).$$

Уровень химического потенциала дырок при этом не изменяется, т.е. уровни химических потенциалов электронов и дырок расщепляются:

$$\mu_n'(T_e) \neq -\varepsilon_g - \mu_p^0(T_0).$$

Однако расщепление уровней химического потенциала в свою очередь означает появление рекомбинации, формирующей в итоге единый уровень химического потенциала электронов и дырок $\mu_n(T_e) = -\varepsilon_g - \mu_p(T_0)$. Действительно, в пространственно однородном, статическом случае в отсутствие внешней генерации уравнения (1) сводятся к равенствам $R_n = 0$ и $R_p = 0$, что имеет место

только при совпадении уровней химических потенциалов носителей. Отметим, что, вообще говоря, изменению подвергаются оба химических потенциала, т.е.

$$\mu_n(T_e) \neq \mu_n^0(T_e), \quad \mu_p(T_0) \neq \mu_p^0(T_0).$$

Другими словами, в результате рекомбинации энергетическая неравновесность приводит к появлению неравновесности концентрационной.

Математический анализ рассматриваемого случая аналогичен рассмотрению состояния термодинамического равновесия в однородном образце, за исключением того, что в модели конкретного механизма рекомбинации должно быть получено соответствующее выражение для ее скорости, которая является функцией не только концентрации носителей, но и температуры дырок (решетки), и температуры электронов. Последнее обусловлено явной температурной зависимостью сечений захвата носителей примесями или друг другом.

В итоге для описания рассматриваемого случая однородного разогрева получаем систему четырех уравнений:

$$R_n(n, p, n_i; T_e, T_0) = 0, \quad R_p(n, p, n_i; T_e, T_0) = 0,$$

а также уравнение для концентрации заполненных примесных уровней n_i и условие электронейтральности образца (т.е. уравнение Пуассона для однородного случая). Одно из этих уравнений является следствием трех других, что обеспечивает разрешимость данной системы уравнений (подчеркнем, что у нас всего три независимых неизвестных, n , p и n_i).

Обратим внимание на тот факт, что, как и в ранее рассмотренной ситуации, выражения для темпов рекомбинации $R_n(n, p, n_i; T_e, T_0)$ и $R_p(n, p, n_i; T_e, T_0)$ в общем случае различаются, т.е. в такой форме записи нет тождества $R_n \equiv R_p$. Однако если с помощью условия электронейтральности или уравнения для примесного уровня (которое в статическом режиме сводится именно к условию $R_n = R_p$) исключить n_i , то мы возвращаемся к тождеству

$$R_n(n, p, n_i; T_e, T_0) \equiv R_p(n, p, n_i; T_e, T_0) \equiv R(n, p; T_e, T_0)$$

в полном соответствии с условием непрерывности полного тока $\text{div } \mathbf{j} = 0$.

Напомним, что температуры носителей мы считаем заданными. В противном случае необходимо добавить уравнение теплового баланса.

Дальнейшие несложные расчеты показывают, что, как и было указано выше, энергетическая неравновесность при наличии температурной зависимости сечений захвата приводит к сдвигу положения динамического равновесия между процессами захвата и обратной термической генерации, тем самым изменяя концентрации носителей. Подробное рассмотрение этого эффекта и анализ условий, при которых он наиболее ярко выражен или пренебрежимо мал, приведены в работе [13].

4. Явления переноса неравновесных носителей

Перейдем к рассмотрению ситуации, когда имеет место протекание тока при наличии в полупроводнике неравновесных носителей. Положим, что токоперенос не сопровождается разогревом образца, т. е. температура постоянна в каждой точке полупроводника и равна своему равновесному значению T_0 . Вначале проанализируем случай, когда неравновесность не является генерационной (т. е. нет внешней объемной генерации неравновесных носителей, например светом). При этом может иметь место поверхностная генерация или инжекция носителей, которые должны быть учтены соответствующими слагаемыми в граничных условиях. Отметим, что корректная формулировка граничных условий при наличии тока весьма важна для правильного описания кинетических явлений в ограниченных полупроводниках. Обсуждению этой проблемы посвящен последний раздел настоящей работы.

В дальнейшем ограничимся анализом только статического токопереноса (концентрация неравновесных носителей не зависит от времени), так как именно в этом режиме наиболее ярко проявляются противоречия традиционного описания кинетических эффектов. Во избежание математических трудностей проведем рассмотрение в линейном приближении, хотя изложенный далее подход справедлив и в общем нелинейном случае. Он может быть использован для построения корректной модели нелинейного токопереноса при компьютерном моделировании.

Всюду далее будем обозначать символом " $\delta \dots$ " отклонения соответствующих величин от их равновесных значений (обозначаемых индексом "0").

Система уравнений, описывающая рассматриваемую ситуацию, имеет следующий вид:

$$\operatorname{div} \mathbf{j}_n = eR_n(n, p, n_t), \quad \operatorname{div} \mathbf{j}_p = -eR_p(n, p, n_t), \quad (6)$$

$$R_n(n, p, n_t) = R_p(n, p, n_t), \quad (7)$$

$$\operatorname{div} \mathbf{E} = 4\pi\rho(n, p, n_t). \quad (8)$$

Приведенная система уравнений является полной, так как для четырех неизвестных n , p , n_t и φ имеем четыре независимых уравнения.

Подчеркнем два обстоятельства. Во-первых, в (8) объемный заряд ρ в общем случае представляет собой функцию не только концентрации носителей n , p , но и концентрации примесных зарядов n_t , которая может изменяться при протекании тока. Во-вторых, еще раз обратим внимание на то, что выражение (7) является уравнением для концентрации электронов на примесном уровне n_t . После исключения последней данное выражение становится тождеством в полном соответствии с уравнением непрерывности заряда $\operatorname{div} \mathbf{j} = 0$.

Определим роль неравновесного заряда захваченных носителей в формировании неравновесного объемного

заряда и критерии пренебрежения величиной n_t . Для этого исключим последнюю из уравнений (6)–(8), выбрав в качестве механизма рекомбинации модель Шокли–Рида [14]. В соответствии с этой моделью скорость захвата электронов примесным уровнем и обратного ему процесса — термической генерации — равны соответственно

$$r_n = \alpha_n n(N_t - n_t), \quad g_{nT} = \alpha_n n_1 n_t, \quad (9)$$

и темп рекомбинации электронов равен $R_n = r_n - g_{nT}$, где N_t — полная концентрация примесных состояний (во избежание недоразумений напомним, что в рассматриваемой модели n_t — концентрация заполненных примесных состояний), α_n — коэффициент захвата электронов, n_1 — параметр, характеризующий примесный уровень и по своему физическому смыслу представляющий концентрацию электронов, которая имела бы место, если бы уровень Ферми полупроводника совпадал с примесным уровнем.

Аналогично для дырок:

$$r_p = \alpha_p n_t p, \quad g_{pT} = \alpha_p p_1(N_t - n_t). \quad (10)$$

Подставляя выражения (9), (10) в (7), в линейном приближении получим для концентрации заполненных примесных состояний следующее соотношение:

$$\delta n_t = a\delta n + b\delta p,$$

где

$$a = \frac{\alpha_n n_1 N_t}{(n_0 + n_1)[\alpha_n(n_0 + n_1) + \alpha_p(p_0 + p_1)]}, \quad (11)$$

$$b = \frac{\alpha_p n_0 N_t}{(n_0 + n_1)[\alpha_n(n_0 + n_1) + \alpha_p(p_0 + p_1)]}. \quad (12)$$

Тогда $R_n = R_p = \delta n/\tau_n + \delta p/\tau_p$, причем $\tau_n/\tau_p = n_0/p_0$. Подчеркнем еще раз, что величины $\tau_{n,p}$ хотя и имеют размерность времени, тем не менее даже в линейном приближении не являются временами жизни носителей в строгом смысле слова.

Как легко заметить, в электронном полупроводнике $\tau_n \gg \tau_p$, в дырочном $\tau_p \gg \tau_n$. Таким образом, в монополярном полупроводнике, в ситуациях, близких к квазилинейности ($\delta n \approx \delta p$), действительно темп рекомбинации определяется характеристикой ("временем жизни") неосновных носителей. В собственном полупроводнике параметры τ_n и τ_p равны. Однако только в случае квазинейтральности они представляют собой по своему физическому смыслу удвоенное время жизни электронно-дырочных пар. Если же квазинейтральность не имеет места, даже в собственном полупроводнике $\delta n \neq \delta p$, то говорить о времени жизни нельзя. Как нетрудно убедиться, проанализировав (11), (12), в случае электронного полупроводника ($n_0 \gg p_0$) и высоких температур (квазиуровни Ферми и электронов, и дырок

далеки от примесного уровня, т.е. донорный уровень полностью ионизован)

$$1 \gg a \gg b, \text{ т.е. } \delta n_t \ll \delta n. \quad (13)$$

Аналогично для дырочного полупроводника ($p_0 \gg n_0$) при высоких температурах выполняется неравенство

$$1 \gg b \gg a, \quad (14)$$

и по-прежнему концентрацией неравновесных захваченных носителей δn_t можно пренебречь.

Ситуация изменится, если температуры низкие или в полупроводнике имеется достаточно высокая концентрация глубоких примесных уровней, расположенных вблизи середины запрещенной зоны, т.е. вблизи уровня Ферми собственного полупроводника в равновесии. В этом случае концентрация δn_t сравнима по порядку величины с концентрацией неравновесных носителей и должна быть учтена в уравнении Пуассона. Другими словами, если квазиуровни Ферми электронов и дырок при протекании тока располагаются заведомо далеко от примесного уровня, то система уравнений (6)–(8) может быть упрощена за счет пренебрежения изменением концентрации связанных зарядов:

$$\operatorname{div} \mathbf{j}_n = eR(n, p), \quad \operatorname{div} \mathbf{j}_p = -eR(n, p), \quad (15)$$

$$\operatorname{div} \mathbf{E} = 4\pi\rho(n, p). \quad (16)$$

Теперь и объемный заряд, и темп рекомбинации зависят только от концентрации неравновесных подвижных носителей δn и δp . Уравнение для примесного уровня (7), представляющее собой в данном случае тождество, естественно, из системы уравнений выпадает.

Также отметим, что если в конкретной задаче окажется, кроме того, $\delta p \ll \delta n_t$, то в уравнении Пуассона (16) в выражении для плотности заряда $\rho(n, p)$ необходимо опустить δp во избежание превышения точности, т.е. считать $\rho = \rho(n)$.

Дальнейшее упрощение системы уравнений может быть осуществлено, если выполняются условия квазинейтральности, т.е. все характерные параметры задачи существенно превышают радиус Дебая (вообще говоря, особенно в нелинейном случае, необходима верификация полученных решений, так как характерные параметры задачи могут являться функциями, например, приложенного поля, а значит, неравновесных концентраций, и в сильных полях могут сравниваться с радиусом Дебая).

В приближении квазинейтральности $\delta\rho = 0$; в силу условий (13), (14) концентрации неравновесных электронов и дырок совпадают: $\delta n = \delta p$; уравнение Пуассона становится излишним, оно может быть в дальнейшем использовано для верификации полученного решения:

$$\operatorname{div} \delta \mathbf{E} \ll 4\pi(\delta p - \delta n)/\epsilon;$$

а токоперенос описывается системой двух уравнений для $\delta n = \delta p$ и φ :

$$\operatorname{div} \mathbf{j}_n = eR(n, p), \quad \operatorname{div} \mathbf{j}_p = -eR(n, p). \quad (17)$$

5. Кинетические явления в условиях генерации носителей

Как мы уже отмечали, предложенное выше рассмотрение применимо, когда отсутствует внешняя генерация носителей (фотогенерация). Для ее учета необходимо последовательно повторить все изложенные выше соображения (стартуя с уравнений (1)), удерживая слабые $g_{n,p}$.

Отметим, что в общем случае темпы генерации электронов и дырок могут быть различны даже в статическом случае (например, имеет место примесное поглощение света и, как результат, генерация носителей только одного типа). В свою очередь в статике и темпы рекомбинации становятся различными, ибо для выполнения условия непрерывности заряда $\operatorname{div} \mathbf{j} = 0$ теперь необходимо выполнение равенства $R_n - R_p + G = 0$, где $G \equiv g_n + g_p$. Последнее равенство по-прежнему представляет собой уравнение кинетики примесных уровней в статическом случае. После исключения концентрации связанных зарядов n_t система уравнений, описывающая статическое протекание тока в условиях постоянной внешней генерации носителей, принимает следующий вид:

$$\frac{1}{e} \operatorname{div} \mathbf{j}_n = \beta_n g_n + \beta_p g_p + R(n, p), \quad (18)$$

$$-\frac{1}{e} \operatorname{div} \mathbf{j}_p = \beta_n g_n + \beta_p g_p + R(n, p), \quad (19)$$

$$\operatorname{div} \mathbf{E} = 4\pi\rho(n, p, g_n - g_p). \quad (20)$$

Подчеркнем, что в отличие от уравнений (1) фигурирующие в (18), (19) "эффективные" темпы генерации

$$g_{\text{eff}} \equiv \beta_n g_p + \beta_p g_n$$

совпадают независимо от величины истинных темпов генерации g_n и g_p . Это обстоятельство обеспечивает тождественное выполнение условия непрерывности полного тока при любых концентрациях неравновесных носителей и произвольных темпах их генерации.

Коэффициенты $\beta_{n,p}$ в уравнениях (18), (19) определяются механизмом рекомбинации этого процесса. Например, для межзонной рекомбинации $\beta_n = \beta_p = 1/2$, для модели рекомбинации Шокли–Рида

$$\beta_n = \frac{\alpha_p(p_0 + p_1)}{\alpha_n(n_0 + n_1) + \alpha_p(p_0 + p_1)},$$

$$\beta_p = \frac{\alpha_n(n_0 + n_1)}{\alpha_n(n_0 + n_1) + \alpha_p(p_0 + p_1)}, \quad (21)$$

но всегда выполняется соотношение $\beta_n + \beta_p = 1$.

Обратим внимание на тот факт, что в статике генерационные члены присутствуют и в уравнении Пуассона, но только в виде разности темпов генерации электронов и дырок. С физической точки зрения это совершенно очевидно — только при различных темпах генерации электронов и дырок возможно изменение концентрации заполненных примесных состояний, т.е. появление

объемного заряда, вызванного непосредственно фотогенерацией.

Для межзонной генерации $g_n \equiv g_p = g$, $\beta_n = \beta_p$ и уравнения токопереноса имеют хорошо известный вид:

$$\frac{1}{e} \operatorname{div} \mathbf{j}_n = g + R(n, p), \quad -\frac{1}{e} \operatorname{div} \mathbf{j}_p = g + R(n, p), \quad (22)$$

$$\operatorname{div} \mathbf{E} = 4\pi\rho(n, p). \quad (23)$$

Мы хотели бы привлечь внимание к следующему обстоятельству. Довольно часто в литературе (см., например, [2]) полагают, что в случае межзонной рекомбинации и генерации всегда имеет место равенство концентраций неравновесных носителей $\delta n = \delta p$. Однако это утверждение никоим образом не следует из уравнений (22), (23). Действительно, если темпы генерации электронов и дырок (как и темпы их рекомбинации) совпадают, то отсюда, вообще говоря, не следует равенство их концентраций (так же, как равенство производных двух функций не обеспечивает равенство их самих). Рождающиеся носители могут быть сразу же сепарированы, например, внешним полем. Для выполнения равенства $\delta n = \delta p$ необходимо выполнение условий квазинейтральности, а никак не наличие только межзонных переходов (т.е. последнее, вообще говоря, является необходимым, но не достаточным условием равенства концентраций неравновесных электронов и дырок).

Для монополярного полупроводника (для определенности, n -типа) $\beta_n \ll \beta_p$ и, если $g_n \approx g_p$, система уравнений (18)–(20) принимает следующий вид:

$$\frac{1}{e} \operatorname{div} \mathbf{j}_n = g_p + R_n(n, p), \quad -\frac{1}{e} \operatorname{div} \mathbf{j}_p = g_p + R_p(n, p), \quad (24)$$

$$\operatorname{div} \mathbf{E} = 4\pi\rho(n, p, g_n - g_p). \quad (25)$$

Подчеркнем, что в этом случае на статический токоперенос непосредственно влияет только генерация неосновных носителей (в то время как генерация основных носителей, если темп отличается от темпа генерации неосновных носителей, проявляет себя опосредовано через уравнение Пуассона).

В качестве иллюстрации использования описанного выше подхода проанализируем случай однородного монополярного полупроводника (для определенности, электронного) при однородной фотогенерации носителей.

Пусть имеет место только генерация основных носителей — электронов ($g_p = 0$), т.е. имеют место только переходы с примесного уровня в зону проводимости и обратно (энергия фотонов превышает энергию ионизации примеси, но меньше ширины запрещенной зоны). Тогда система уравнений, описывающая данную ситуацию, может быть записана в следующем виде:

$$g_n + R_n(n, n_t) = 0, \quad \operatorname{div} \mathbf{E} = 4\pi\rho(n, n_t, g_n). \quad (26)$$

Отметим, что, как это обычно принято в такого рода ситуациях, мы пренебрегли наличием неосновных носителей ($p = 0$, $R_p = 0$).

В силу однородности рассматриваемой задачи уравнение Пуассона вырождается в условие нейтральности образца $\delta n = -\delta n_t$ (причем нейтральность имеет место при произвольных размерах полупроводника, в отличие от приближения квазинейтральности).

При высокой температуре ($n_1 \gg N_t$) решение системы уравнений (26) имеет следующий вид:

$$\delta n = \frac{g_n}{\alpha_n n_1}, \quad \delta n_t = -\delta n, \quad \delta p = 0. \quad (27)$$

Если не пренебрегать неосновными носителями, то, естественно, мы приходим к такому же результату, т.е. генерация основных носителей практически не влияет на концентрацию неосновных носителей, $\delta p \ll \delta n$.

Рассмотрим теперь противоположный предельный случай, т.е. генерацию только неосновных носителей ($g_n = 0$). Отметим, что, хотя практически реализация такой ситуации маловероятна (так как для генерации неосновных носителей требуется энергия фотонов, заведомо достаточная для генерации основных носителей), с методической точки зрения этот случай достаточно интересен и дает ключ к пониманию процессов, протекающих при наличии генерации различной интенсивности носителей обоих типов. Как нетрудно убедиться, решение системы уравнений для этого предельного случая приводит к следующему результату:

$$\delta n = \frac{n_1 g_p}{\alpha_p n_0 N_t}, \quad \delta p = \delta n, \quad \delta n_t \ll \delta n. \quad (28)$$

Отметим, что первое из выражений (28) можно представить как

$$\delta n = \frac{n_1}{N_t} \frac{n_1}{n_0} \frac{\alpha_n}{\alpha_p} \delta n_{(27)}, \quad (29)$$

где $\delta n_{(27)}$ — концентрация неравновесных основных фотоносителей в рассмотренной ранее ситуации генерации только основных носителей, т.е. результат (27). Так как $n_1 \gg N_t$, а из определения n_1 следует, что при высоких температурах $n_1 \gg n_0$, очевидно, что генерация неосновных носителей изменяет концентрацию основных гораздо более эффективно, чем их прямая генерация. Обратим также внимание и на тот факт, что, хотя $\delta n_t \ll \delta n$, тем не менее $\delta n_t \gg \delta n_{(27)}$.

Таким образом, если в полупроводнике одновременно генерируются и основные, и неосновные носители, то генерация последних оказывает более существенное влияние как на заполнение примесных состояний, так и на концентрацию основных носителей, и, как следствие, на протекание тока или формирование фотоэдс.

6. Генерационно-рекомбинационные процессы в неоднородных температурных полях

Проанализируем установление стационарного режима токопереноса в неоднородном температурном поле $T = T_0 + \delta T(x)$. Для упрощения выкладок предположим, что неоднородность температуры мала, т.е. $\delta T(x) \ll T_0$ и градиент температуры постоянен $\nabla T = \text{const}$. При описании этой простейшей, казалось бы, термоэлектрической задачи тем не менее возникает ряд проблем, наиболее очевидных в биполярном полупроводнике. Основная проблема кроется в форме записи рекомбинационного члена. В линейном приближении последний может быть представлен в виде $R(\delta n, \delta p)$. Однако при этом возникает вопрос — от какой величины должна отсчитываться концентрация неравновесных носителей в неоднородном температурном поле: $n_0(T_0)$, $n_0[T(x)]$ или $n_0[T(x)]$. Впервые ответ на этот вопрос был дан в работе [9] на основе рассмотрения поэтапного установления стационарного режима в температурном поле. По аналогии с ранее рассмотренным примером становления равновесия в неоднородно-легированном полупроводнике на первом этапе в каждом мысленно изолированном слое устанавливается равновесная концентрация носителей, соответствующая данной локальной температуре $n = n_0(T_0) + \delta n'[\delta T(x)]$. Это изменение концентрации за счет неоднородности температуры можно представить как появление неоднородности химического потенциала, но при постоянной температуре. На втором этапе под действием градиента химического потенциала происходит перетекание электронов и дырок. Итоговое распределение носителей на втором этапе обеспечивает появление встроенного термоэлектрического поля и равенство нулю градиента электрохимического потенциала $\nabla(\mu - e\varphi) = 0$. Естественно, результирующее распределение носителей тока на втором этапе должно обеспечивать единство уровня химического потенциала электронов и дырок ($\mu_p = -\varepsilon_g - \mu_n$), что и является одним из условий нахождения поправок к концентрациям. Именно найденное на втором этапе распределение концентрации носителей должно рассматриваться как равновесное в формулах для рекомбинации. На третьем этапе градиент температуры выступает в роли внешней силы, и с учетом этого находится действительное распределение концентрации носителей в полупроводнике в неоднородном температурном поле.

Отметим, что описанная выше процедура оказывается весьма удачной с математической точки зрения, поскольку на каждом этапе (подчеркнем, что в отличие от метода последовательных приближений в рассматриваемом случае все поправки одного порядка) получаются линейные дифференциальные уравнения с постоянными коэффициентами.

Недостатком описанного метода, помимо его некоторой громоздкости, является невозможность использовать

его, если температурное поле не задано, а должно быть найдено из самосогласованного решения транспортной задачи. Выходом в этом случае является возвращение к рассмотрению статистики переходов носителей между зонами и примесными уровнями с учетом температурной зависимости рекомбинационных характеристик.

Нетрудно убедиться, проделав несложные выкладки, аналогичные приведенным в разд. 4, что в общем случае выражение для рекомбинации при статическом токопереносе в неоднородном температурном поле в линейном приближении при любом механизме рекомбинации имеет следующий вид:

$$R_n = R_p = \frac{\delta n}{\tau_n} + \frac{\delta p}{\tau_p} + \gamma \delta T, \quad (30)$$

где параметр γ равен

$$\gamma = \frac{2}{\tau_n} \frac{n_i}{p_0} \frac{\partial n_i}{\partial T} = \frac{2}{\tau_p} \frac{n_i}{n_0} \frac{\partial n_i}{\partial T}. \quad (31)$$

Здесь n_i — значение концентрации носителей в собственном полупроводнике при температуре T_0 .

Поучительно отметить, что темп генерации (рекомбинации) в полупроводнике в неоднородном температурном поле, вообще говоря, оказывается отличным от нуля даже в отсутствие тока и фотовозбуждения. Таким образом, в тех задачах, когда этим эффектом пренебречь нельзя (например, при описании фотоакустического эффекта, подробнее см. [15]), необходимо рассматривать не только баланс энергии, но и транспортные уравнения (1). Как легко заметить из (31), это особенно актуально для собственных полупроводников с достаточно узкой запрещенной зоной и, как следствие, высокой концентрацией собственных носителей.

7. Граничные условия

Мы бы хотели обратить внимание читателя еще на один момент, очень часто служащий источником неточностей и ошибок при описании транспортных процессов и не нашедший должного, на наш взгляд, отражения в литературе. Приведенные выше уравнения, будучи дифференциальными уравнениями, требуют задания граничных условий.

В общем случае на поверхности образца имеют место дополнительные, специфические механизмы рекомбинации носителей, которые должны быть корректно учтены именно в граничных условиях. Поскольку статистически поверхностная рекомбинация аналогична объемной рекомбинации, все изложенные выше в данной работе соображения должны быть применены и при написании граничных условий.

Кроме того, необходимо обратить внимание на возможность существования и движения каждого типа носителей в каждой из контактирующих сред. В качестве примера укажем структуру "металл–дырочный полупроводник–металл". В отсутствие рекомбинации

(как объемной, так и поверхностной), поскольку в металле дырочный ток равен нулю, токоперенос через дырочный полупроводник будет осуществляться только неосновными носителями — электронами. Ясно, что сопротивление образца в этом случае окажется существенно выше, чем можно было ожидать [9].

К сожалению, очень часто используемые даже в широко известной литературе граничные условия не являются корректными. Приведем лишь один пример, подтверждающий сказанное. При рассмотрении транспортных задач в полупроводниковых структурах весьма широко используются граничные условия следующего вида (например, для электронов, независимо от того, являются ли они основными или неосновными носителями) [2]:

$$j_n|_I = s\delta n|_I,$$

где s — скорость поверхностной рекомбинации на границе раздела двух проводящих сред I . Однако данное условие по своему физическому смыслу означает, что носители не переносятся через контакт, а генерируются или рекомбинируют на нем, т.е. весь ток в цепи формируется исключительно механизмами поверхностной рекомбинации. Как правило, в реальных структурах это не так.

Другим примером является игнорирование поверхностной рекомбинации в транспортных задачах без должной оценки корректности такого приближения. Такой подход может привести к неправильным результатам, ибо влияние поверхностной рекомбинации на транспортные процессы имеет место уже в линейном по полю приближении [1]. Весьма эффективным средством решения указанных проблем представляется изложенная в [10] методика построения корректных граничных условий на основе введения феноменологических констант, характеризующих перенос носителей через границу раздела двух сред. При необходимости эти константы должны быть рассчитаны с использованием микроскопической теории конкретного контакта.

8. Заключение

Приведенный выше анализ показывает, что в рамках единых физических представлений возможно описать явления переноса в полупроводниках и полупроводниковых структурах, не детализируя конкретные механизмы генерационно-рекомбинационных процессов. Показано, что наличие температурных полей и рассогласование температур носителей тока и фононов может существенно повлиять на рекомбинационные процессы при изучении явлений переноса даже в линейном по внешним воздействиям приближении.

Выяснен физический смысл широко используемого приближения квазинейтральности и изучены следствия, к которым это приближение может привести.

Авторы благодарят CONACyT, Мéxico, за частичную поддержку при выполнении этой работы.

Список литературы

- [1] Ю.Г. Гуревич, Г.Н. Логвинов, Г. Эспехо, О.Ю. Титов, А. Мериуц, ФТП, **34**, 783 (2000).
- [2] В.Л. Бонч-Бруевич, С.Г. Калашников. *Физика полупроводников* (М., Наука, 1990).
- [3] R.A. Smith. *Semiconductors* (University Press, Cambridge, 1961).
- [4] П.И. Баранский, В.П. Клочков, И.В. Потыкевич. *Полупроводниковая электроника. Справочник* (Киев, Наук. думка, 1975).
- [5] И. Аут, Д. Генцов, К. Герман. *Фотоэлектрические явления* (М., Мир, 1980).
- [6] С.М. Рывкин. *Фотоэлектрические явления в полупроводниках* (М., ГИФМЛ, 1963).
- [7] D.A. Neamen. *Semiconductor Physics and Devices: Basic Principles*, ed. by Richard D. Irwin (Boston, 1992).
- [8] С. Зи *Физика полупроводниковых приборов* (М., Мир, 1984).
- [9] Yu.G. Gurevich, G.N. Logvinov, O.I. Lubimov, O.Yu. Titov. *Phys. Rev. B*, **51**, 6999 (1995).
- [10] Yu.G. Gurevich. *J. Thermoelectricity*, N 2, 5 (1997).
- [11] I.M. Паренко, С.Л. Королюк, В.М. Кошкін, С.С. Москалюк. *Науковий вісник Чернівецького університету, Фізика*. № 29, 45 (1998).
- [12] В.П. Силин, А.А. Рухадзе. *Электромагнитные свойства плазмы и плазмодобных сред* (М., Госатомиздат, 1961).
- [13] Yu.G. Gurevich, I.N. Volovichev. *Phys. Rev. B*, **60**, 7715 (1999).
- [14] W. Shockley, W. T. Read, *Phys. Rev.*, **87**, 835 (1952).
- [15] I.N. Volovichev, L. Villegas, G. González de la Cruz, Yu.G. Gurevich. *J. Appl. Phys.* (in press).

Редактор Т.А. Полянская

Generation and recombination in semiconductors

I.N. Volovichev*†, Yu.G. Gurevich†

* Institute for Radiophysics and Electronics, National Academy of Sciences of Ukraine, 310085 Kharkov, Ukraine

† Departamento de Física, CINVESTAV-IPN, 07000 México, D.F., México

Abstract A unified technical approach to study of transport phenomena in semiconductors is suggested. A variety of models of recombination processes, used in research of transport phenomena and the establishment of equilibrium in semiconductor structures, have been analyzed. A new expression describing steady-state recombination in arbitrary temperature fields is presented. Manifestations of the quasineutrality condition in thermodynamic equilibrium and in transport phenomena have been observed.