1946 г. Т XXVIII, вып. 4

УСПЕХИ ФИЗИЧЕСКИХ НАУК

ЭЛЕКТРОННЫЕ ПРОЦЕССЫ В РЕАЛЬНЫХ КРИСТАЛЛАХ

Ф. Ф. Волькенштейн

§ 1. Введение § 2. Идеальные и реальные кристаллы § 3. Теория свободных электронов Зоммерфельда § 4. Зональная теория § 5. Проводники и изоляторы с точки зрения зональной теории § 6. Электронные уровни в реальной решётке § 7. Электропроводность в неметаллических кристаллах § 8. Влияние примесей и температуры на электропроводность полупроводников § 9. Деформация энергетических зон под влиянием внешнего электрического поля § 10. Ионизирующее действие электрического поля § 11. Фотоэлектрические процессы

§ 1. ВВЕДЕНИЕ

Все макроскопические свойства кристаллов могут быть разделены на два класса. К одному классу принадлежат все те свойства, которые определяются периодической структурой решётки и для которых учёт дефектов, неизбежно присутствующих во всякой реальной решётке, играет роль незначительной поправки. К другому классу относятся все те свойства, которые, наоборот, определяются именно локальными нарушениями в периодической структуре решётки, для которых, следовательно, учёт дефектов имеет принципиальное значение. Свойства, принадлежащие к этому второму классу, обычно называют структурно-чувствительными свойствами.

К структурно-чувствительным свойствам принадлежат оптические и фотоэлектрические свойства щёлочно-галоидных кристаллов, электропроводность полупроводников и другие. К этому же классу свойств относится активность кристаллических катализаторов, а также оптические свойства фосфоров.

Все эти свойства чрезвычайно чувствительны ко всякого рода примесям, содержащимся в кристалле. Слово «примесь», однако, не следует понимать здесь слишком буквально как наличие в решётке чужеродных атомов. «Примесью» является всякое нарушение периодической структуры решётки. В частности, в решётках, изображающихся химической формулой M_nR_r, где M — символ металла, а R — символ металлоида, под «примесью» можно понимать

всякое отклонение от стехиометрического соотношения. В формуле $M_m R_r$, такие отклонения могут быть отражены, если индексы m и r считать не целыми, а дробными числами, хотя и близкими к целым.

Так, под влиянием примесей в миллионы раз изменяется электропроводность полупроводников, изменяется окраска кристаллов (т. е. интенсивность и положение полос поглощения). Изменяется (увеличивается или уменьшается) каталитическая активность катализаторов. В изменении всех этих разнообразных свойств наблюдается известный параллелизм.

Можно думать, что все эти разнообразные свойства кристаллов в конечном счёте имеют одно и то же происхождение. Действительно, кристалл в целом представляет собой некую единую электрическую систему. В этой системе разыгрываются своеобразные электронные процессы. Ими обусловливаются как устойчивые, так и структурно-чувствительные свойства кристаллов. Два фактора, борющиеся друг с другом, управляют этими процессами: фактор порядка и фактор беспорядка. Первый из них предопределяет устойчивые свойства кристалла. Вторым — обусловливаются структурно-чувствительные свойства.

Таким образом, при теоретической интерпретации структурно-чувствительных свойств мы заведомо должны исходить из представлений о реальной кристаллической решётке, т. е. такой решётке, строго-периодическая структура которой нарушена. Если базироваться на идеальной решётке, как это делают в большинстве вопросов теории металлов, то все структурно-чувствительные свойства кристаллов заведомо останутся за рамками теории.

Теория электрических свойств полупроводников, с одной стороны, и теория оптических и фотоэлектрических свойств щёлочно-галоидных кристаллов, с другой стороны, представляют собой как бы две различные главы одной и той же теории реального кристалла вообще. Хотя щёлочно-галоидные кристаллы по своим макроскопическим свойствам существенно отличаются от полупроводников, тем не менее между этими двумя классами кристаллов нет принципиального различия. В обоих случаях мы имеем кристаллы типа $M_m R_r$ с нарушенной стехиометрией.

Как щёлочно-галоидным кристаллам, так и кристаллам полупроводников соответствует одна и та же схема энергетических уровней, и различие между ними сводится лишь к различию в масштабе, принятом на этой схеме. Все оптические и фотоэлектрические свойства щёлочно-галоидных кристаллов должны повторяться в полупроводниках, но должны быть при этом перенесены из ультрафиолетовой и видимой частей спектра в красную и инфракрасную область, плохо доступную экспериментальному исследованию. С другой стороны, все электрические свойства полупроводников, теоретически говоря, должны наблюдаться и в щёлочно-галоидных кристаллах, но при

значительно более высоких температурах, лежащих за пределами экспериментальных возможностей.

Теория каталитических свойств кристаллических (неметаллических) катализаторов представляет собой как бы ещё одну главу всё той же общей теории реального кристалла. Действительно, во всех этих случаях мы имеем дело по существу с одним и тем же объектом, рассматриваемым лишь под различными углами зрения.

§ 2. ИДЕАЛЬНЫЕ И РЕАЛЬНЫЕ КРИСТАЛЛЫ

Реальный кристалл отличается от идеального наличием дефектов. Под словом «дефект» мы будем понимать всякое нарушение в периодической структуре решётки. Дефекты, присутствующие в реальной решётке, могут быть двух сортов: дефекты макроскопического и дефекты микроскопического характера.

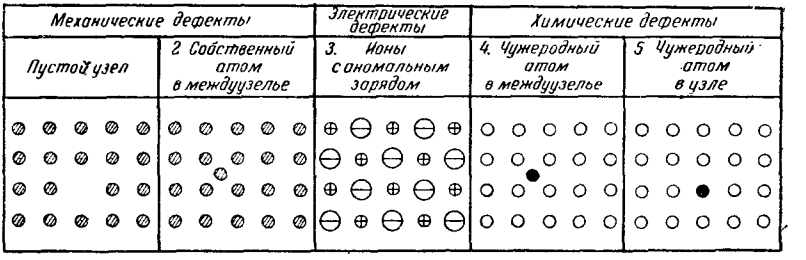

Рис. 1. Основные типы микродефектов.

К макроскопическим дефектам относятся трещины, коллоидные частицы постороннего вещества, наблюдающиеся, например, в окрашенных щёлочно-галоидных кристаллах, всякого рода макроскопические вкрапления и пр.

Что касается микроскопических дефектов, то отметим здесь следующие основные типы дефектов этого сорта:

1) пустой узел, образованный в результате удаления из идеальной решётки атома или иона;

2) собственный атом или ион решётки, расположенный в междуузелье;

3) ион в гетерополярной решётке, находящийся в нормальном положении, но несущий на себе аномальный заряд;

4) чужеродный атом, помещённый в междуузелье;

5) чужеродный атом, помещённый в узел, т. е. замещающий собственный атом решётки.

На рис. 1 схематически изображены эти пять типов микродефектов. Первые два из них можно назвать механическими дефек-

тами. Такие дефекты в случае гомеополярной решётки не изменяют её химического состава, а в случае гетерополярной решётки могут приводить к тому или иному нарушению стехиометрического соотношения. Дефекты третьего типа могут быть названы электрическими дефектами. Они также не влияют на химическую формулу решётки. Наконец, два последних типа микродефектов можно назвать химическими дефектами. Это есть «примесь» в узком смысле этого слова, искажающая химический состав кристалла.

Химические дефекты вводятся в решётку извне. Их возникновение обусловлено той или иной обработкой образца. Механические дефекты могут возникать не только путём введения их извне, но и в результате нагревания решётки. Наконец, электрические дефекты не только могут быть введены извне или созданы теплом; они могут возникать в решётке также под влиянием освещения.

Рассмотрим особо гетерополярные решётки типа M_mR_r, обладающие механическими микродефектами. К этому типу решёток относятся почти все полупроводники.

Структурными элементами такой решётки являются положительные ионы металла с зарядом $+p$ и отрицательные ионы металлоида с зарядом $-q$, где

$$\frac{q}{p} = \frac{m}{r}.$$

Здесь p и q — абсолютная величина заряда, выраженная в единицах заряда электрона. Мы имеем в нашей решётке два сорта узлов: металлические узлы, в которых помещаются ионы металла, и металлоидные узлы, в которых сидят ионы металлоида.

Если все металлические и металлоидные узлы решётки замещены соответствующими ионами и если при этом не существует ионов, расположенных вне этих узлов, то наша решётка является идеальной решёткой. В такой решётке стехиометрическое соотношение заведомо соблюдено.

Практически, однако, мы всегда имеем дело с решётками, в которых стехиометрическое соотношение в той или иной степени нарушено. От характера и степени этого нарушения резко зависят все каталитические, электрические и оптические свойства кристалла. Стехиометрические нарушения свидетельствуют о наличии дефектов в решётке.

Стехиометрические нарушения в кристалле могут быть двоякого рода: мы можем иметь решётку со стехиометрическим избытком металла или со стехиометрическим избытком металлоида. Избыток металла может быть осуществлён за счёт введения в решётку дополнительных ионов металла или за счёт удаления из неё ионов металлоида. Избыток металлоида может быть осуществлён путём введения ионов металлоида или удаления ионов металла.

Таким образом, в случае катализатора или полупроводника мы имеем дело с реальной гетерополярной решёткой, которая отличается от идеальной следующими чертами. В такой решётке не все узлы замещены ионами: в ней имеется некоторое число пустых металлических и металлоидных узлов. Кроме того, в такой решётке не все ионы помещаются в узлах: в ней имеется некоторое число м е ж д у у з е л ь н ы х металлических и металлоидных ионов.

В такой реальной гетерополярной решётке стехиометрическое соотношение может быть соблюдено или же может быть нарушено в ту или иную сторону и в той или иной степени в зависимости от того, каково соотношение между числом дефектов разных сортов.

По той роли, которую эти дефекты играют в решётке, они могут быть разбиты на две группы. Междуузельные металлические ионы и пустые металлоидные узлы могут быть названы э л е к т р о п о л о ж и т е л ь н ы м и дефектами. Для электрона, введённого в нашу решётку, эти дефекты служат центрами притяжения. Междуузельные металлоидные ионы и пустые металлические узлы могут быть названы э л е к т р о о т р и ц а т е л ь н ы м и дефектами. По отношению к электрону они играют роль центров отталкивания.

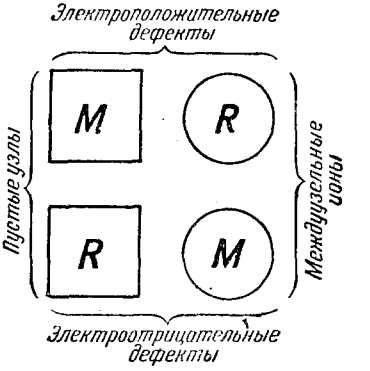

Рис. 2. Классификация дефектов (по их роли в решётке и по их физической природе).

Такая классификация дефектов (по их физической природе и по той роли, которую они играют в решётке) поясняется схемой, изображённой на рис. 2.

Общее число электроположительных и электроотрицательных дефектов, содержащихся в решётке, характеризует степень отклонения нашей реальной решётки от идеального состояния. Разность между числом электроположительных и электроотрицательных дефектов определяет степень нарушения стехиометрического соотношения. Преобладание электроположительных дефектов свидетельствует о стехиометрическом избытке металла. Стехиометрический избыток металлоида характеризуется, наоборот, преобладанием электроотрицательных дефектов.

Степень отклонения реальной решётки от идеальной, а также степень отклонения от стехиометрического соотношения представляют собой две н е з а в и с и м ы е характеристики. Таким образом, всякая реальная гетерополярная решётка для своего описания требует введения д в у х параметров. Этими параметрами определяются все основные свойства решётки. Варьируя эти параметры, мы можем получить различные типы температурных кривых электропроводности, различные типы внутреннего фотоэффекта (положительный или отри-

цательный фотоэффект или фотоэлектрически неактивное поглощение) и пр. [1].

Значения этих параметров определяются двумя факторами.

Прежде всего, они определяются биографией решётки, т. е. зависят от того, каким способом был изготовлен образец и каким воздействиям он подвергался в течение всей своей предшествующей жизни.

Во-вторых, они зависят от температуры. Действительно, при нагревании ионы решётки диссоциируют из своих узлов, образуя междуузельные ионы и пустые узлы. Температура, таким образом, вызывает в решётке дополнительный беспорядок. При повышении температуры степень отклонения реальной решётки от идеальной возрастает (но степень отклонения от стехиометрического соотношения остаётся при этом неизменной).

Таким образом, беспорядок в решётке по своему происхождению может быть врождённым (необратимым) или тепловым (обратимым).

В некоторых случаях тепловым беспорядком можно пренебречь по сравнению с врождённым беспорядком, иначе говоря, можно пренебречь дефектами, вызванными температурой, по сравнению с дефектами, запасёнными в решётке с самого начала и присутствующими в ней, так сказать, в замороженном состоянии. Это имеет место в достаточно испорченных решётках при неслишком высоких температурах. С этим случаем мы имеем дело в теории полупроводников. В этом случае число дефектов в решётке можно считать практически не зависящим от температуры. В качестве дефектов мы имеем здесь или пустые узлы или междуузельные ионы, в зависимости от природы решётки и той обработки, которой она была подвергнута. Прокаливая образец в вакууме или в той или иной атмосфере при том или ином давлении, мы можем изменять характер дефектов и их концентрацию в решётке.

В других случаях, наоборот, тепловой беспорядок преобладает над врождённым беспорядком. В этом случае можно пренебречь числом замороженных дефектов по сравнению с числом тепловых дефектов. Это имеет место, повидимому, в случае щёлочно-галоидных кристаллов, о чём свидетельствует хорошо известный экспериментальный факт: экспоненциальное возрастание числа центров окрашивания с температурой. В этом случае решётка содержит в себе как междуузельные ионы, так и пустые узлы. Стехиометрическое соотношение в этом случае остаётся практически ненарушенным.

Мы видели, что идеальная решётка превращается в реальную в результате удаления ионов из узлов или введения ионов в междуузелье. Заметим, что эта операция обязательно должна сопровождаться введением или удалением из решётки электронов. Этот второй процесс должен сопутствовать первому для того, чтобы электрический заряд, образующийся вследствие введения или удаления ионов,

был бы скомпенсирован, т. е. для того, чтобы была сохранена электрическая нейтральность решётки. Вместо того, чтобы говорить об удалении электрона, мы можем говорить о введении «дырки» и наоборот. Наличие дырки означает отсутствие электрона. Такой способ выражаться оказывается весьма удобным.

Создание в решётке электроположительного дефекта требует введения некоторого числа электронов. Создание электроотрицательного дефекта должно сопровождаться введением некоторого числа дырок. Эти электроны и дырки, присутствующие во всякой реальной решётке для соблюдения электрического баланса, вообще говоря, не обязаны рекомбинировать друг с другом. В некоторых решётках такая рекомбинация энергетически выгодна, в других — невыгодна, т. е. приводит к повышению общей энергии системы. Эти электроны и дырки играют весьма важную роль. Они составляют тот резерв, за счёт которого при надлежащих условиях возникает электропроводность решётки. Ими обусловливаются также оптические свойства решётки.

Общий запас электронов и дырок зависит от того, насколько сильно наша решётка отличается от идеальной. Преобладание же электронов или дырок в этом общем запасе выражено тем более резко, чем сильнее нарушено стехиометрическое соотношение. В решётках со стехиометрическим избытком металла мы имеем избыток электронов; в решётках со стехиометрическим избытком металлоида мы имеем недостаток электронов, т. е. избыток дырок.

Компенсирующие электроны и дырки, введённые в решётку и рассаженные по её ионам, приводят к образованию аномальных ионов, т. е. ионов с аномальным зарядом. Это могут быть, например, металлические ионы с зарядом $+p-1$ или металлоидные ионы с зарядом $-q+1$. Таким образом, гетерополярная решётка, обладающая механическими дефектами, вообще говоря, должна содержать в себе также и электрические дефекты, обеспечивающие её электрическую нейтральность.

Аномальные ионы с одноимёнными дополнительными зарядами отталкиваются друг от друга, а аномальные ионы с разноимёнными дополнительными зарядами притягиваются друг к другу. Здесь правильнее было бы сказать, что отталкиваются или притягиваются не аномальные ионы, а аномальные состояния ионов. Кроме того, аномальные ионы с отрицательным дополнительным зарядом притягиваются к электроположительным дефектам, а ионы с положительным дополнительным зарядом—к электроотрицательным дефектам. Таким образом, электроположительные дефекты служат ловушками для электронов, а электроотрицательные дефекты играют роль ловушек по отношению к дыркам.

В нормальном состоянии все электроны и дырки локализованы около соответствующих дефектов. Таким образом, весь запас электронов и дырок, содержащихся в решётке, законсервирован. Воздей-

ствуя на решётку теми или иными внешними агентами (нагревание, освещение и др.), можно вывести электроны и дырки из сферы действия дефектов. Этими освобождёнными электронами и дырками обеспечивается электронная и дырочная проводимость решётки.

§ 3. ТЕОРИЯ СВОБОДНЫХ ЭЛЕКТРОНОВ ЗОММЕРФЕЛЬДА

Кристаллическая решётка представляет собой систему взаимодействующих электронов, движущихся в поле положительных ядер, которые в первом приближении можно считать закреплёнными в некоторых жёстко зафиксированных положениях равновесия. Задача заключается в определении поведения такой системы. Эта задача не может быть решена точно. Обычно, при её решении пользуются теми или иными упрощениями.

Первое упрощение, которое мы сделаем, будет заключаться в замене взаимодействия электронов некоторым эффективным силовым полем, в котором каждый из электронов будем считать движущимся независимо от других (метод «самосогласованного» поля). Таким путём наша

Рис. 3. Потенциальная энергия электрона в решётке.

многоэлектронная задача сводится к задаче одноэлектронной. Каждому электрону может быть приписано своё индивидуальное состояние и своя индивидуальная энергия. Поле, в котором движется электрон, имеет периодический характер. Оно схематически изображено на рис. 3, *а*. Поведение электрона в таком периодическом поле рассматривается в теории Блоха [2].

Второе упрощение состоит в замене периодического потенциала некоторым средним его значением, постоянным по всему объёму кристалла. Иначе говоря, потенциальная кривая, изображённая на рис. 3, *а*, заменяется апроксимированной кривой, изображённой на рис. 3, *б*. Из такой апроксимации исходит теория Зоммерфельда [3].

Мы рассмотрим сперва нашу задачу в рамках теории Зоммерфельда, а затем посмотрим, как изменяются наши результаты, при учёте периодического потенциала.

В теории Зоммерфельда мы имеем дело, по существу, с той же моделью кристалла, из которой исходила классическая теория металлов, развитая Друде [4]. Электроны в кристалле трактуются как электронный газ, запертый в потенциальном ящике с плоским дном.

Однако, теория Зоммерфельда отличается от теории Друде следующими двумя существенными чертами.

Первая отличительная черта заключается в том, что в теории Зоммерфельда поведение каждого отдельного электрона подчиняется законам квантовой механики, в то время как в теории Друде оно описывалось методами классической механики.

Эта реформа имеет своим следствием то, что возможные для электрона значения импульса образуют теперь не непрерывный, а дискретный ряд. Электрон в кристалле может обладать только таким импульсом $\mathbf{p}$, составляющие p_1, p_2, p_3 которого удовлетворяют условиям

$$p_i = \frac{h}{2L} n_i, \qquad (1)$$

где $i = 1, 2, 3$. Здесь L — длина кристалла (кристалл мы представляем себе в виде куба), а n_i — любое целое число:

$$n_i = 0, \pm 1, \pm 2, \ldots$$

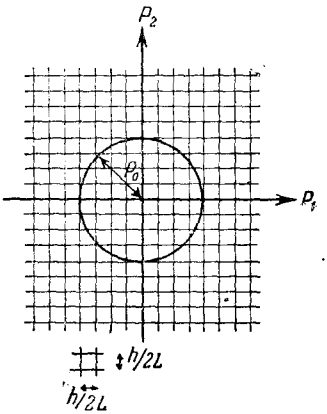

Рис. 4. Пространство импульсов.

Пространство импульсов, представленное на рис. 4, распадается на ячейки объёма $\Delta\gamma = h^3/8L^3$. Электрон может находиться лишь в тех точках этого пространства, которые соответствуют узлам «решётки», изображённой на рис. 4. Вместе с тем делается дискретным также энергетический спектр. Для энергии W электрона мы имеем согласно (1):

$$W = \frac{\mathbf{p}^2}{2m} = \frac{h^2}{8mL^2}(n_1^2 + n_2^2 + n_3^2), \qquad (2)$$

где m — масса электрона. Целочисленные переменные n_1, n_2, n_3 играют роль квантовых чисел. Зависимость p_i и W от n_i изображена на рис. 5, *а* и 5, *б*. На вертикальной оси рис. 5, *б* нанесена система энергетических уровней электрона.

Вторая черта, отличающая теорию Зоммерфельда от теории Друде, заключается в том, что при описании поведения всей системы

электронов в целом классическая статистика Максвелла-Больцмана заменяется квантовой статистикой Ферми-Дирака.

Статистика Ферми-Дирака, в отличие от статистики Максвелла-Больцмана, имеет дело с тождественными частицами, т. е. с частицами принципиально неразличимыми. Это значит, что различные состояния системы, отличающиеся перестановкой частиц, рассматриваются как одно и то же состояние.

Кроме того, частицы, с которыми имеет дело статистика Ферми-Дирака, подчиняются принципу Паули. В случае невзаимодействующих частиц принцип Паули сводится к требованию, чтобы все частицы нашей системы находились в различных индивидуальных состояниях. Это значит, что в одной и той же точке пространства импульсов, изображённого на рис. 4, не может помещаться одновременно более одного электрона (при игнорировании спина), подобно тому, как не могут находиться несколько электронов в одной и той же точке обычного координатного пространства.

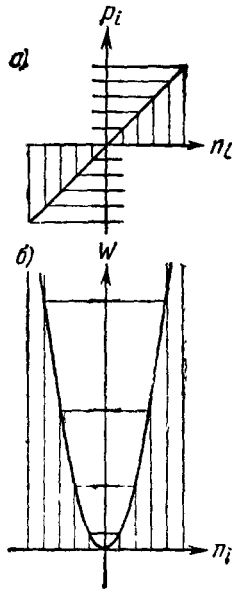

Рис. 5. Возможные значения импульса и энергии для электрона в решетке (теория Зоммерфельда).

Учёт обоих этих обстоятельств (тождественность частиц и принцип Паули) отражается на окончательных формулах. Заметим, однако, что все формулы квантовой статистики практически превращаются в классические формулы в следующих предельных случаях:

а) когда частицы обладают достаточно малой массой;

б) когда температура достаточно высока;

в) когда концентрация частиц достаточно мала.

В случае электронного газа в металле переход от статистики Ферми к статистике Максвелла происходит при очень высоких температурах, лежащих значительно выше тех температур, с которыми имеет дело экспериментатор. В случае электронного газа в полупроводнике, благодаря малой плотности этого газа, такой переход совершается, наоборот, при очень низких температурах, лежащих ниже температур, представляющих практический интерес. Поэтому электронный газ в металле должен описываться формулами Ферми, в то время как электронный газ в полупроводнике может быть описан формулами Максвелла.

На рис. 6 изображена полная энергия W электронного газа, как функция температуры T. Кривая AB соответствует квантовой статистике, а прямая OB — классической статистике. Наклон кривой $W = W(T)$ характеризует собой теплоёмкость C электронного газа.

Действительно, по определению:

$$C = \frac{dW}{dT}.$$

Точка T_0 на рис. 6 делит всю шкалу температур на две области:

1) при $T \gg T_0$ квантовая статистика практически совпадает с классической (невырожденный газ);

2) при $T \ll T_0$ квантовая статистика существенно отличается от классической (вырожденный газ).

Рассматривая рис. 6, мы можем сделать два вывода, являющихся характерными следствиями применения к электронному газу статистики Ферми:

а) в области достаточно низких температур (в области вырождения) теплоёмкость электронного газа практически равна нулю;

б) при абсолютном нуле температуры энергия электронного газа отлична от нуля.

При абсолютном нуле размещение электронов по ячейкам фазового пространства (рис. 4) соответствует наиболее плотной упаковке. Электроны заполняют сферу, изображённую на рис. 4, радиус p_0 которой определяется из условия

$$\frac{{}^4/_3\pi p_0^3}{\Delta\gamma} = \frac{N}{2},$$

где N — общее число электронов (на единицу объёма кристалла); двойка в правой части равенства появляется вследствие учёта спина. При нагревании плотная упаковка электронов разрыхляется: внутри сферы радиуса p_0 появляются свободные места (дырки) за счёт электронов, выброшенных за пределы этой сферы. При этом, однако, симметрия в распределении электронов по импульсам, характерная для абсолютного нуля, сохраняется. Это значит, что ток отсутствует. Появление тока (при наложении электрического поля) свидетельствует о нарушении симметрии. Таким образом, температура и электрическое поле представляют собой факторы, вызывающие перераспределение электронов в фазовом пространстве.

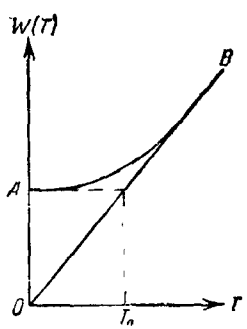

Рис. 6. Полная энергия электронного газа как функция температуры.

OB — статистика Максвелла-Больцмана, AB — статистика Ферми-Дирака.

§ 4. ЗОНАЛЬНАЯ ТЕОРИЯ

Теория Зоммерфельда обладает рядом недостатков. В частности, оставаясь в рамках теории Зоммерфельда, нельзя построить оптики металлов. Действительно, зоммерфельдовский свободный электронный газ не может поглощать света: его взаимодействие со световыми

квантами сводится к Комптон-эффекту. Кроме того, с точки зрения теории Зоммерфельда нельзя объяснить различия между металлическими и неметаллическими кристаллами.

Эти недостатки обусловлены теми грубыми приближениями, которые лежат в основе зоммерфельдовской модели. Они исчезают при переходе к теории, учитывающей периодический потенциал решётки.

Учёт периодического потенциала отражается прежде всего на структуре энергетического спектра. Энергетические уровни, изображённые на рис. 5, *б*, стягиваются в отдельные группы, образуя так называемые энергетические зоны, разделённые более или менее широкими запрещёнными участками.

Структура зон и их расположение в спектре могут быть выяснены, если решать задачу, исходя из изолированных атомов, как это делает Блох[2], и рассматривать решётку как результат сближения таких первоначально изолированных атомов. В методе Блоха внутренние электроны атома в совокупности с ядром рассматриваются как положительно заряженный атомный остов, в поле которого движутся валентные электроны, в первом приближении не взаимодействующие друг с другом. Поведение электрона внутри отдельного атома считается известным. Задача состоит в том, чтобы определить, что происходит с электроном и с его системой энергетических уровней под влиянием остальных атомов решётки. Действие этих последних рассматривается как слабое возмущение. Строго говоря, этот метод приемлем лишь в случае решёток, построенных из одновалентных атомов.

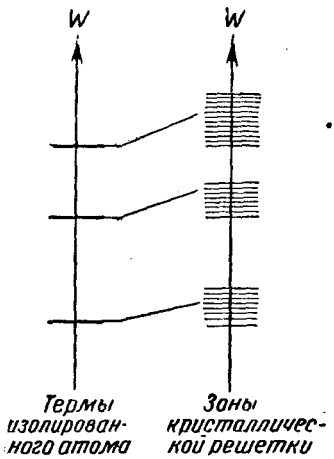

Рис. 7. Энергетический спектр решётки.

В теории Блоха каждая зона является результатом расщепления соответствующего терма изолированного атома. Это расщепление обусловлено взаимодействием каждого данного атома с остальными атомами решётки. Действительно, каждый атом решётки находится в возмущающем поле своих соседей. Каждый терм порождает свою зону, как это изображено на рис. 7. Таким образом, на энергетическом спектре решётки отражается индивидуальность тех атомов, из которых построена решётка.

Рассмотрим *s*-терм изолированного атома. Этот терм отличается от прочих тем, что он не вырожден. При соединении атомов в решётку *s*-терм превращается в *s*-зону. В случае кубической решётки структура этой зоны, как показал Блох, определяется

формулой

$$W = W_0^s + \alpha^s + 2\beta^s\left(\cos\frac{2\pi a}{L}n_1 + \cos\frac{2\pi a}{L}n_2 + \cos\frac{2\pi a}{L}n_3\right), \qquad (3)$$

где a — постоянная решётки; n_1, n_2, n_3 — квантовые числа. Эти числа могут принимать следующие значения:

$$n_i = 0, \pm 1, \pm 2, \ldots, \pm\frac{L}{2a}. \qquad (4)$$

Мы получаем целую систему уровней соответственно всем возможным значениям квантовых чисел n_1, n_2, n_3. Совокупность этих уровней образует зону. Тройка квантовых чисел n_1, n_2, n_3 может трактоваться как номер уровня внутри зоны. Все уровни располагаются между верхней и нижней границами зоны.

Верхняя граница зоны: $W_{\max} = W_0^s + \alpha^s + 6\beta^s$

Середина зоны: $W_0 = W_0^s + \alpha^s$

Нижняя граница зоны: $W_{\min} = W_0^s + \alpha^s - 6\beta^s$

Ширина зоны: $\Delta W = W_{\max} - W_{\min} = 12\beta^s$

Мы видим, что параметр α^s в (3) характеризует сдвиг середины зоны относительно исходного терма W_0^s. Параметр β^s характеризует ширину зоны. Значение этого параметра зависит от того, насколько сильно перекрываются волновые функции двух соседних атомов. Чем выше лежит исходный терм W_0^s и чем теснее расположены атомы в решётке, тем сильнее это перекрытие, тем больше β^s и тем шире зона.

Каждой тройке квантовых чисел n_1, n_2, n_3 соответствует определённое состояние электронов, т. е. определённая волновая функция, описывающая поведение электрона. Эти волновые функции обладают следующим общим свойством. Модуль каждой волновой функции является периодической функцией координат с периодом, равным периоду решётки. Это значит, что электрон с равной вероятностью может находиться около любого атома решётки. Иначе говоря, все атомы решётки участвуют в игре на равных правах. С другой стороны, это означает, что наш электрон полностью коллективизирован. Он принадлежит всем атомам решётки в одинаковой степени.

Из (4) видно, что зона содержит в себе конечное число состояний, равное числу атомов в решётке. Так как каждое состояние может быть занято только двумя электронами с противоположными спинами, то, следовательно, зона обладает определённой ёмкостью по отношению к электронам. Она может вместить в себя лишь ограниченное число электронов.

Электрону, находящемуся в состоянии n_1, n_2, n_3, может быть приписан импульс $\mathbf{p}$ (n_1, n_2, n_3). Это есть среднее значение импульса, соответствующее групповой скорости волнового пакета. Импульс $\mathbf{p}$

как функция квантовых чисел n_1, n_2, n_3 может быть вычислен, если известна энергия W как функция n_1, n_2, n_3. Связь между $\mathbf{p}$ и W даётся формулой Джонеса-Зинера

$$\mathbf{p} = \frac{Lm}{h} \operatorname{grad}_{\mathbf{n}} W. \tag{5}$$

Значок $\mathbf{n}$ при градиенте обозначает, что здесь имеется в виду градиент не в обычном координатном пространстве x, y, z, а в пространстве квантовых чисел n_1, n_2, n_3, так что составляющие вектора $\mathbf{p}$ имеют вид:

$$p_i = \frac{Lm}{h} \partial W / \partial n_i,$$

где $i = 1, 2, 3$. Подставляя сюда (3), будем иметь:

$$p_i = \frac{4\pi am}{h} \beta^s \sin \frac{2\pi a}{L} n_i. \tag{6}$$

Формулы (6) и (3) заменяют собой соответствующие формулы (1) и (2) зоммерфельдовской теории. Зависимость p_i и W от n_i изображена на рис. 8 (ср. рис. 5).

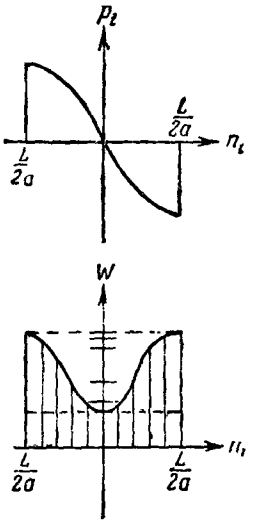

Рис. 8. Возможные значения импульса и энергии для электрона в решётке (теория Блоха).

Сравнивая (6) и (3), мы видим, что одной и той же энергии W соответствуют два состояния с противоположно направленными импульсами. Это следует из того, что энергия W является чётной функцией от n_1, n_2, n_3, в то время как импульсы p_i являются нечётными функциями этих параметров. Таким образом, на одном и том же энергетическом уровне могут помещаться два электрона, движущихся в противоположные стороны.

Заметим, что около верхней границы зоны и около её нижней границы косинусы в (3) можно разложить в ряд и ограничиться первыми членами разложения. Тогда выражение (3) примет такой вид:

$$W = W_0^s + a^s \pm 6\beta^s \pm$$
$$\pm \frac{4\pi^2 a^2}{L^2} \beta^s (n_1^2 + n_2^2 + n_3^2),$$

где верхний знак соответствует разложению около верхней границы, а нижний знак — разложению около нижней границы зоны. Сравнивая эту формулу с формулой (2) для свободных электронов, мы замечаем, что электрон, находящийся на дне зоны или поднятый под потолок зоны, может трактоваться как свободный электрон зоммерфельдовской теории с той лишь разницей, что теперь ему следует приписать не-

которую эффективную массу m^* вместо истинной массы m

$$m^* = \mp \frac{h^2}{32\pi^2 a^2 \beta^s}.$$

Абсолютная величина этой массы тем меньше, чем шире зона, причём эта масса положительна на дне зоны и отрицательна под потолком зоны. Это значит, что под потолком зоны электрон ведёт себя весьма своеобразным образом: он ускоряется в сторону, противоположную действующей на него силе.

Рассмотрим теперь зоны, соответствующие вырожденным термам. В качестве примера такого терма рассмотрим p-терм. Этот терм трёхкратно вырожден. Его можно рассматривать как совокупность трёх термов, совпадающих друг с другом. p-терм порождает p-зону, которую можно представлять себе состоящей из трёх зон, наложенных друг на друга. В случае кубической решётки структура этих трёх зон, как показал Вильсон[6], определяется формулами:

$$\left\{ \begin{array}{l} W^1 = W(n_1, n_2, n_3), \\ W^2 = W(n_2, n_1, n_3), \\ W^3 = W(n_3, n_2, n_1), \end{array} \right\} \tag{7}$$

где

$$W(n_1, n_2, n_3) = \\ = W_0^p + \alpha^p + 2\beta^p \cos\frac{2\pi a}{L} n_1 + 2\gamma^p \left(\cos\frac{2\pi a}{L} n_2 + \cos\frac{2\pi a}{L} n_3\right),$$

Квантовые числа n_1, n_2, n_3 попрежнему принимают значения (4). Мы видим, что все три зоны (7) имеют общие границы:

$$W^1_{\max} = W^2_{\max} = W^3_{\max} = W_0^p + \alpha^p + 2\beta^p + 4\gamma^p,$$
$$W^1_{\min} = W^2_{\min} = W^3_{\min} = W_0^p + \alpha^p - 2\beta^p - 4\gamma^p.$$

Зоны, порождаемые более высокими термами (d-терм, f-терм и т. д.), имеют более сложную стуктуру[7]. Число состояний, содержащихся внутри такой сложной зоны, всегда равно числу атомов в решётке, умноженному на кратность вырождения соответствующего терма. Модули волновых функций всегда остаются периодическими. Для любой зоны сохраняется общая формула Джонеса-Зинера (5), выражающая связь между энергией W и импульсом $\mathbf{p}$ электрона.

Отдельные зоны в энергетическом спектре решётки, вообще говоря, разделены запрещёнными участками. Если, однако, два соседних терма в энергетическом спектре изолированного атома расположены достаточно близко друг к другу, а порождаемые ими зоны достаточно широки, то запрещённый участок между зонами может оказаться достаточно узким. В частности, он сводится к нулю, когда две соседние зоны соприкасаются или перекрываются друг с другом.

В этом случае соседние зоны нельзя рассматривать раздельно. Они как бы вступают во взаимодействие друг с другом, и их структура искажается.

До сих пор мы рассматривали нашу задачу, как одноэлектронную. При строгом учёте взаимодействия электронов наша многоэлектронная задача уже не может быть сведена к одноэлектронной. В этом случае приписывание отдельному электрону своей системы энергетических уровней теряет смысл. Теперь можно говорить лишь об энергетическом спектре всей системы электронов в целом.

Точное решение многоэлектронной задачи оказывается невозможным. В теории металлов взаимодействие электронов обычно учитывают приближённо, не выходя за рамки одноэлектронной задачи. Это достигается, как мы уже отмечали, путём прибавления к периодическому потенциалу решётки некоторого дополнительного так называемого самосогласованного потенциала[8]. Самосогласованный потенциал представляет собой энергию взаимодействия данного электрона с размазанным зарядом всех электронов нашей системы, в том числе и со своим собственным размазанным зарядом. Таким образом, взаимодействие заменяется введением некоторого эффективного поля, в котором электроны движутся независимо друг от друга. Самосогласованный потенциал отличается важным свойством: он обладает той же периодичностью, что и потенциал решётки.

Метод самосогласованного потенциала является весьма несовершенным методом. Мы можем ожидать от него удовлетворительных результатов лишь в случае гомеополярной решётки, построенной из одновалентных атомов. При переходе к гетерополярной решётке применимость этого метода делается сомнительной.

С точки зрения теории Блоха гетерополярную решётку следует представлять себе состоящей из системы положительных ионов и семейства коллективизированных электронов, составленного из валентных электронов как металлических, так и металлоидных атомов. Электронный газ, с которым мы имеем здесь дело, обладает большой плотностью: на каждую кристаллическую ячейку приходится по нескольку электронов. Каждый электрон движется в поле металлических и металлоидных положительных ионов и в самосогласованном поле электронов.

Возможен, однако, иной подход к задаче[1], более совершенный с точки зрения учёта взаимодействия электронов. Гетерополярную решётку можно представлять себе построенной из положительных металлических и отрицательных металлоидных ионов. В этом случае семейство коллективизированных электронов, движущихся в поле этих ионов, составляется из случайных электронов, внесённых в решётку извне или сорванных с отдельных ионов. В этом случае плотность электронного газа ничтожно мала по сравнению с плотностью электронного газа в методе Блоха. Поэтому взаимодействием электронов с полным правом можно совсем пренебречь. По этой же причине

можно пренебречь искажающим влиянием аномальных ионов, присутствующих в решётке. Действительно, число таких аномальных ионов, отдавших свои электроны в общественное пользование, не может превышать числа коллективизированных электронов. Поле решётки, в котором движутся коллективизированные электроны, сохраняет свойство периодичности. Поэтому, при нашем подходе к задаче энергетический спектр решётки сохраняет зональную структуру, характерную для теории Блоха.

§ 5. ПРОВОДНИКИ И ИЗОЛЯТОРЫ С ТОЧКИ ЗРЕНИЯ ЗОНАЛЬНОЙ ТЕОРИИ

Для того чтобы решётка обладала проводимостью, необходимо соблюдение двух условий:

а) решётка должна содержать свободные электроны, т. е. такие электроны, которые могут беспрепятственно передвигаться по всему объёму кристалла;

б) внешнее электрическое поле должно обладать способностью в известной степени регулировать беспорядочное движение электронов, вызывая необходимую асимметрию в распределении электронов по скоростям.

В классической теории первое из этих двух условий служило критерием того, имеем ли мы дело с проводником или изолятором. В квантовой теории таким критерием служит второе условие.

Различие между свободными и связанными электронами, в классическом понимании этих терминов, может быть пояснено с помощью рис. 3, *а*. На этом рисунке изображено периодическое поле, в котором движется электрон: потенциальные ямы чередуются с потенциальными барьерами. Полная энергия W электрона изображается на рис. 3, *а* горизонтальным уровнем. Этот уровень может лежать над вершинами потенциальных барьеров (уровень AA') или под ними (уровень BB'). В первом случае мы имеем дело со свободным электроном. Такой электрон беспрепятственно переходит от атома к атому поверх разделяющих их барьеров. Во втором случае мы имеем дело со связанным электроном. С классической точки зрения такой электрон вынужден навсегда оставаться внутри своего атома; действительно, чтобы перейти в соседний атом, электрон должен пройти через запрещённую область (заштрихованную на рис. 3,*а*), внутри которой его полная энергия меньше потенциальной, т. е. кинетическая энергия отрицательна. Атомы оказываются разделёнными стенками, совершенно непроницаемыми для классического электрона.

В квантовой теории резкое различие между свободными и связанными электронами, в классическом понимании этих терминов, стирается, поскольку квантовая механика позволяет электрону переходить

из атома в атом не только поверх барьера, но и сквозь барьер. Если в классической теории вероятность такого перехода могла быть равна единице (свободный электрон) или нулю (связанный электрон), то теперь эта вероятность может принимать все промежуточные значения. Она тем больше, чем меньше площадь, заштрихованная на рис. 3, *а*. Таким образом, с точки зрения квантовой теории все электроны кристалла являются «свободными» в том смысле, что все они могут с той или иной степенью лёгкости передвигаться по кристаллу.

Различие между проводником и изолятором с квантовой точки зрения обусловлено зональной структурой энергетического спектра решётки.

Если число коллективизированных электронов в нашей решётке таково, что при размещении их по уровням зоны (по два электрона на каждое состояние) они заполняют зону целиком и, если при этом наша заполненная зона отделена запрещённым участком от выше

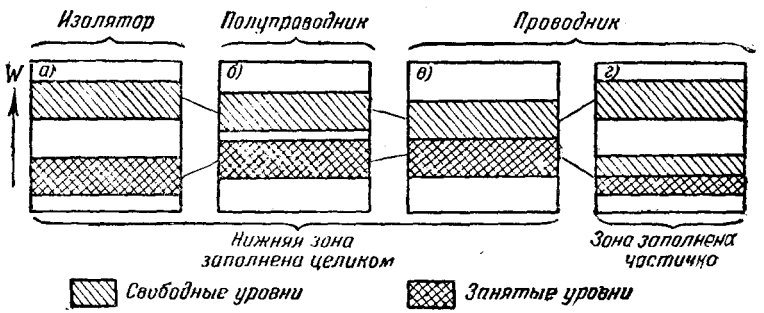

Рис. 9. Различие между проводником и изолятором (с точки зрения зональной теории).

лежащей пустой зоны (см. рис. 9, *а*), то в этом случае решётка представляет собой изолятор. Действительно, плотная упаковка электронов по уровням зоны соответствует симметричному распределению по скоростям. Электрону с импульсом **p** всегда может быть приведён в соответствие другой электрон с импульсом — **p**. Перераспределение электронов по импульсам в пределах данной зоны оказывается невозможным. Единственно, что можно делать, это переставлять электроны один на место другого, но таким путём нельзя получить желаемой асимметрии в распределении по импульсам. Внешнее поле оказывается не в силах как бы то ни было влиять на движение электронов, находящихся в данной зоне, и, следовательно, решётка оказывается нечувствительной к воздействию внешнего поля.

Представим себе теперь, что общее число коллективизированных электронов меньше, чем удвоенное число состояний внутри зоны. В этом случае электроны заполняют зону не целиком. При максимально

плотной упаковке занятыми оказываются лишь нижние уровни зоны. По соседству с занятыми уровнями имеются пустые уровни (см. рис. 9, *г*). Пересаживая электроны с одних уровней на другие в пределах данной зоны, можно получить асимметричное распределение по импульсам, свидетельствующее о том, что число электронов, движущихся в одном направлении, преобладает над числом электронов, движущихся в противоположном направлении. Роль внешнего поля сводится именно к такому пересаживанию электронов, нарушающему симметрию в распределении по импульсам. Наша решётка является в этом случае проводником.

Представим себе, наконец, что зона заполнена электронами до отказа, но при этом соприкасается или перекрывается с вышележащей пустой зоной (см. рис. 9,*в*). В этом случае решётка опять-таки будет представлять собой проводник. Действительно, верхняя пустая зона является в этом случае продолжением нижней заполненной зоны. Перераспределение электронов по импульсам может быть осуществлено за счёт уровней этой верхней зоны. Такие две соприкасающиеся или перекрывающиеся зоны можно рассматривать как одну зону, заполненную электронами лишь частично.

Рассмотрим решётку изолятора. В результате внешних воздействий в такой решётке может быть искусственно создана некоторая проводимость.

С точки зрения классической теории это может быть осуществлено в результате перевода связанных электронов в свободное состояние. Для этого надо поднять электрон с уровней BB' на рис. 3, *а* на уровни AA', расположенные поверх потенциальных барьеров. Этот процесс, как это видно из рис. 3, *а*, требует затраты энергии.

С точки зрения квантовой теории это может быть осуществлено за счёт перевода электронов из нижней зоны в верхнюю зону. Этот процесс опять-таки связан с затратой энергии, как это видно на рис. 9,*а*.

Нижнюю заполненную зону изолятора мы будем называть нормальной зоной. Верхнюю пустую зону зоной проводимости. Электроны, попавшие в зону проводимости, в полной мере приобретают способность переносить ток. Дырки, образовавшиеся в нормальной зоне в результате удаления из неё электронов, также служат носителями тока. В решётке изолятора, в которой в результате внешних воздействий часть электронов переброшена из нормальной зоны в зону проводимости, ток складывается из электронной и дырочной компонент.

Внешние факторы, вызывающие проводимость в изоляторе, мы будем называть ионизирующими агентами.

Примером такого агента может служить свет. В этом случае перевод электрона в зону проводимости происходит за счёт энергии светового кванта.

Другим примером может служить температура. При нагревании максимально плотная упаковка электронов внутри нормальной зоны,

соответствующая абсолютному нулю температуры, начинает разрыхляться: электроны выбрасываются из неё в зону проводимости. Число электронов, поднятых в зону проводимости тем больше, чем выше температура, а при заданной температуре — чем у́же запрещённый участок между зонами. Кристаллы с достаточно широким запрещённым участком (рис. 9, *а*) остаются идеальными изоляторами даже при максимальных экспериментально достижимых температурах. Кристаллы с достаточно узким запрещённым участком (рис. 9,*б*) обнаруживают заметную проводимость уже при комнатных температурах. Такие кристаллы обычно называют полупроводниками.

Мы видим, что с точки зрения зональной теории различие между изоляторами, полупроводниками и проводниками (металлами) имеет чисто количественное происхождение. Изолятор можно рассматривать как один крайний случай полупроводника, в то время как металл представляет собой другой крайний случай.

§ 6. ЭЛЕКТРОННЫЕ УРОВНИ В РЕАЛЬНОЙ РЕШЁТКЕ

Переход от идеальной решётки к реальной имеет своё отражение в структуре энергетического спектра.

На систему энергетических зон, характерную для идеальной решётки, накладывается система дискретных, так называемых локальных уровней. Происхождение энергетических зон обусловлено периодической структурой решётки. Происхождение локальных уровней связано с локальными нарушениями этой периодичности. Всякий дефект в решётке, представляя собой такое нарушение, приводит к возникновению дискретного уровня. Расположение энергетических зон отражает в себе природу тех атомов, из которых построена решётка. Расположение локальных уровней в спектре отражает природу дефектов, содержащихся в решётке.

Модули волновых функций, соответствующих уровням зоны, являются периодическими функциями координат с периодом решётки. Это значит, что электрон, сидящий внутри зоны, размазан по всему объёму кристалла, или, иначе говоря, свободно странствует по кристаллу. Локальные уровни, наоборот, характеризуются волновыми функциями, обладающими резко выраженными максимумами около мест повреждения и быстро спадающими по мере удаления от повреждения. Это значит, что электрон, попавший на такой уровень, оказывается локализованным в более или менее узкой области пространства.

Локальные уровни могут быть донорами или акцепторами. Донорами называются такие локальные уровни, которые в нормальном состоянии заняты электронами. Акцепторами называются свободные локальные уровни. Локальные уровни, попадая в запрещённый участок между нормальной зоной и зоной проводимости изолятора, играют важную роль в электронных процессах. Доноры служат поставщиками

электронов. С этих уровней могут черпаться электроны проводимости. Акцепторы играют роль ловушек для электронов: они захватывают и связывают электроны.

Электронными переходами между энергетическими зонами обусловливаются устойчивые свойства кристалла. Переходами, в которых принимают участие локальные уровни, обусловливаются структурно-чувствительные свойства. Энергетический спектр кристаллической решётки является ключом к пониманию всех основных её свойств. Из анализа энергетического спектра вытекают, как следствия, все электрические, фотоэлектрические и оптические свойства решётки.

Рассмотрим различные типы реальных решёток с точки зрения их энергетического спектра.

Простейшим примером реальной решётки может служить решётка, ограниченная с одной стороны. Такая решётка была рассмотрена Таммом[9]. Наличие ограничивающей поверхности уже само по себе является нарушением периодической структуры решётки. Действительно, идеальная решётка, строго говоря, должна быть безграничной. Обрыв решётки приводит к возникновению так называемых поверхностных уровней. Соответствующие волновые функции имеют максимум на поверхности кристалла и быстро спадают по мере удаления от поверхности в глубь кристалла. Это значит, что электроны способны прилипать к поверхности кристалла. Заметим, что любые макроскопические дефекты в решётке характеризуются поверхностными уровнями Тамма. Действительно, всякого рода трещины, поры, макроскопические включения в кристалле свидетельствуют об обрыве решётки.

В качестве другого примера реальной решётки рассмотрим гомеополярную решётку с пустыми узлами. Будем предполагать, что возмущающий потенциал дефекта быстро спадает с расстоянием, так что наличие дефекта практически чувствуют на себе лишь непосредственно соседние к дефекту атомы. В этом случае локальный уровень вырожден, причём кратность вырождения равна числу атомов, непосредственно окружающих дефект. Положение локальных уровней в спектре может быть различным, в зависимости от природы решётки. Если локальные уровни попадают в запрещённый участок между зонами, то соответствующие волновые функции обладают тем более резко выраженными максимумами, чем больше расстояние между локальным уровнем и зоной. По мере уменьшения этого расстояния волновые функции делаются всё более и более расплывчатыми. Когда это расстояние обращается в нуль, локальный уровень как таковой исчезает. Однако, число состояний внутри зоны соответствующим образом увеличивается. При этом все волновые функции (вернее, модули волновых функций) имеют периодический характер.

В качестве третьего примера реальной решётки рассмотрим гомеополярную решётку с вкрапленными в неё чужеродными атомами, размещёнными в междуузельях. В этом случае, как показывает расчёт,

в энергетическом спектре наряду с локальными уровнями, обусловленными самим фактом нарушения периодичности, появляются ещё независимые от них собственные уровни чужеродных атомов. Их положение в спектре определяется как природой чужеродных атомов, так и природой основных атомов решётки.

Рассмотрим наконец гетерополярную решётку, в которой в качестве дефектов фигурируют или пустые узлы двух сортов (металлические и металлоидные) или междуузельные ионы также двух сортов (металлические и металлоидные). Частным случаем таких решёток являются решётки с нарушенным стехиометрическим соотношением. Посмотрим, как выглядит энергетический спектр таких решёток. Положим для простоты, что мы имеем дело с решётками, построенными из однозарядных ионов (решётки щёлочно-галоидных кристаллов). Кроме того, будем считать, что концентрация дефектов практически не меняется с температурой.

В энергетическом спектре решётки мы будем иметь две системы локальных уровней соответственно двум сортам дефектов, присутствующих в решётке: уровни, обусловленные электроположительными дефектами, и уровни, обусловленные электроотрицательными дефектами. Эти уровни располагаются в запрещённом участке между нормальной зоной и зоной проводимости. Пусть на единицу объёма кристалла мы имеем:

$$\begin{cases} N_1 \text{ — число электроположительных уровней (дефектов),} \\ N_2 \text{ — число электроотрицательных уровней (дефектов),} \end{cases}$$

$$\begin{cases} \text{при стехиометрическом избытке металла } N_1 > N_2, \\ \text{при соблюдённом стехиометрическом соотношении } N_1 = N_2, \\ \text{при стехиометрическом избытке металлоида } N_1 < N_2. \end{cases}$$

Для обеспечения электрической нейтральности мы должны внести в нашу решётку N_1 электронов и N_2 дырок. В нормальном состоянии эти компенсирующие электроны локализованы около электроположительных дефектов, а компенсирующие дырки связаны с электроотрицательными дефектами. Таким образом, электроположительные дефекты являются донорами, а электроотрицательные дефекты — акцепторами.

Представим себе, что электроположительные уровни W_1 располагаются в энергетическом спектре ниже электроотрицательных уровней W_2, как это изображено на рис. 10, *а*. На языке нашей энергетической схемы рекомбинация электрона с дыркой означает переход электрона с уровня W_1 на уровень W_2. На рис. 10, *а* такой переход указан стрелкой. Этот переход, как мы видим, связан с затратой энергии и, следовательно, энергетически невыгоден.

Предположим теперь, что электроположительные уровни W_1 расположены выше электроотрицательных уровней W_2. Этот случай изображён на рис. 10, *б*. Рекомбинация электрона с дыркой, как и в предыдущем случае, выражается переходом электрона с уровня W_1 на уровень W_2. Однако, теперь такие переходы (изображённые стрел-

ками на рис. 10, *б*) энергетически выгодны. В решётке, предоставленной самой себе и изолированной от воздействия каких-либо возбуждающих агентов, электроны покинут уровни W_1 и перейдут на уровни W_2, в результате чего энергетические схемы, изображённые на рис. 10, *б*, должны быть заменены энергетическими схемами, изображёнными на рис. 10, *в*. При соблюдённом стехиометрическом соотношении все электроположительные уровни W_1 сделаются акцепторами, а все электроотрицательные уровни W_2 превратятся в доноров. При стехиометрическом избытке металла на уровнях W_1 останутся лишь те электроны, для которых нехватило места на уровнях W_2. При

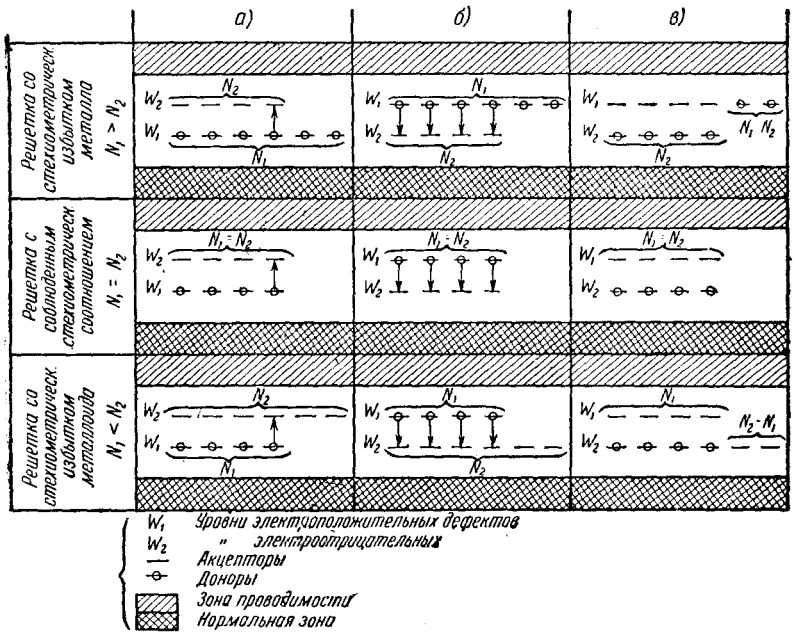

Рис. 10. Схема электронных уровней (гетерополярная решётка с механическими дефектами двух сортов).

стехиометрическом избытке металлоида среди уровней W_2 останутся свободными только те, для заполнения которых нехватило электронов.

Рис. 10, *а* и рис. 10, *в* соответствуют двум различным типам решёток. В решётках первого типа (рис. 10, *а*) электроположительные и электроотрицательные дефекты в нормальном состоянии нейтрализованы компенсирующими электронами и дырками. В таких решётках ионная проводимость отсутствует. Действительно, в таких решётках дефекты могут перемещаться лишь вместе с прилипшими к ним аномальными ионами, т. е. лишь в качестве электрически нейтральных образований. В решётках второго типа (рис. 10, *в*) нейтрализована

лишь та часть дефектов, которая определяет собой отклонение от стехиометрического соотношения. Электроположительными и электроотрицательными дефектами, лишёнными нейтрализующих оболочек, обеспечивается ионная проводимость решётки.

Энергетические схемы, изображённые на рис. 10 и построенные здесь применительно к щёлочно-галоидным кристаллам, могут быть перенесены на общий случай гетерополярных решёток, изображающихся формулой $M_m R_r$, т. е. на случай полупроводников. Полупроводники описываются той же энергетической схемой, что и щёлочно-галоидные кристаллы. Различие между ними сводится лишь к масштабу, принятому на этой схеме. Если энергетический спектр щёлочно-галоидного кристалла сжать, сблизив все уровни, то мы получим спектр полупроводника.

В заключение отметим ещё один тип локальных уровней, обусловленный не механическими, а электрическими дефектами в решётке. Представим себе гетерополярную решётку, лишённую механических дефектов. Предположим, что такая решётка содержит в себе некоторое количество свободных электронов и дырок. Наличие таких электронов и дырок свидетельствует о существовании в решётке ионов с аномальным зарядом. Вокруг таких аномальных ионов решётка естественно оказывается деформированной. Перемещение аномального состояния с одного иона на другой должно сопровождаться соответствующим перемещением всей области деформации. Таким образом, свободный электрон (дырка) в процессе своего движения сам создаёт себе ловушки, в которых он может застрять, поскольку деформация решётки, возникающая вокруг свободного электрона (дырки), является фактором, локализующим его. Локальные уровни такого происхождения были рассмотрены Френкелем[10], а также Ландау[11]. В реальных гетерополярных решётках, обладающих механическими дефектами (пустые узлы, междуузельные ионы), локальные уровни этого сорта играют, повидимому, второстепенную роль.

§ 7. ЭЛЕКТРОПРОВОДНОСТЬ В НЕМЕТАЛЛИЧЕСКИХ КРИСТАЛЛАХ

Электропроводность кристаллической решётки по своей природе может быть ионной и электронной. Необходимым и достаточным критерием ионной проводимости служит эффект Фарадея. Достаточным, но не необходимым критерием электронной проводимости является эффект Холла. В зависимости от того, какие частицы служат носителями тока, ионная проводимость может быть: катионной, анионной или смешанной. Электронная проводимость может быть: чисто электронной, чисто дырочной или смешанной.

Строго говоря, в электропроводности смешаны в той или иной пропорции все составляющие. Однако, соотношение между различными составляющими различно в зависимости от условий. Внешними воздействиями на решётку можно увеличивать роль той или иной из соста-

вляющих. Обычно, одна какая-нибудь из них существенно преобладает, так что остальными можно пренебречь. Заметим, что подвижность ионов очень мала по сравнению с подвижностью электронов. Поэтому преобладание электронной составляющей над ионной наступает даже тогда, когда число свободных электронов ещё очень мало по сравнению с числом свободных ионов.

В полупроводниках электронная проводимость существенно преобладает над ионной. Изоляторы (например, щёлочно-галоидные кристаллы), наоборот, обладают в обычных условиях ионной проводимостью. Это связано с тем, что энергия, необходимая для перевода электрона из связанного состояния в свободное, в случае полупроводников значительно меньше, чем в случае изоляторов. В области сильных полей, по мере возрастания поля, соотношение между ионной и электронной составляющими в изоляторе быстро изменяется в пользу последней. Это происходит потому, что ионная составляющая практически не зависит от поля; иначе говоря, ионная составляющая, вплоть до весьма высоких полей, подчиняется закону Ома. Для электронной же составляющей закон Ома, имеющий место в слабых полях, при переходе к сильным полям сменяется законом Пуля, т. е. экспоненциальным возрастанием проводимости с полем. Это поясняется

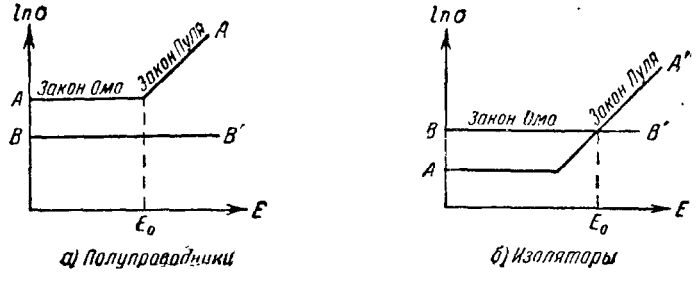

Рис. 11. Электронная и ионная составляющие проводимости как функции поля.

AA' — электронная составляющая, BB' — ионная составляющая.

рис. 11, на котором кривая AA' схематически изображает электронную составляющую проводимости, а кривая BB' — ионную составляющую как функцию поля. Естественно, что переход от закона Ома к закону Пуля у изоляторов происходит при значительно более сильных полях, нежели у полупроводников (точка E_0 на рис. 11). Причём, в случае изолятора переход к закону Пуля свидетельствует о перемене природы проводимости.

Рассмотрим подробнее механизм ионной проводимости.

До тех пор, пока решётка остаётся идеальной, ионная проводимость невозможна. Действительно, единственно возможный способ перемещения ионов в идеальной решётке, не нарушающий её идеаль-

ности, заключается в обмене местами между одноимёнными ионами. Однако, такой обмен местами не может дать тока.

Для того, чтобы сделать возможной ионную проводимость, необходимо предварительно нарушить идеальность решётки, выведя, например, некоторое количество ионов из узлов в междуузелья. Такого рода нарушения осуществляются, как мы знаем, температурой. Междуузельные ионы способны перемещаться внутри решётки и служат, таким образом, носителями тока. С другой стороны, пустые узлы также способны странствовать по решётке. Их так же можно рассматривать как носителей тока. Соответственно этому мы различаем два механизма ионной проводимости:

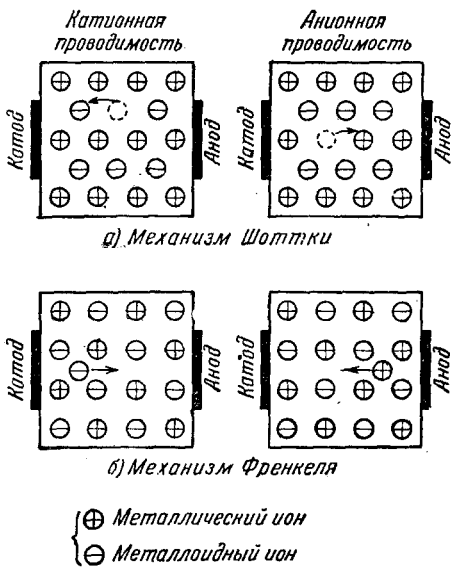

Рис. 12. Механизмы ионной проводимости.

1. Механизм Шоттки [12] (носителями тока являются пустые узлы).

2. Механизм Френкеля [13] (носителями тока являются междуузельные ионы).

Эти два сорта носителей, вообще говоря, обладают разными подвижностями. Число носителей обоих сортов в реальной решётке не обязано быть одинаковым. Действительно, реальная решётка уже с самого начала, т. е. при абсолютном нуле температуры, может обладать некоторым запасом пустых узлов или междуузельных ионов.

Строго говоря, когда мы имеем дело с ионной проводимостью, одновременно действуют оба механизма: механизм Френкеля и механизм Шоттки. Однако, в одних случаях руководящая роль принадлежит механизму Френкеля, в других, наоборот — механизму Шоттки. Эти два механизма схематически иллюстрированы на рис. 12.

Обратимся теперь к электронной проводимости.

Рассмотрим сперва идеальную решётку типа $M_m R_r$, построенную из металлических ионов с зарядом $+p$ и металлоидных ионов с зарядом $-q$. В такой решётке электронная проводимость невозможна. Чтобы сделать возможной электронную проводимость, решётку надо предварительно возбудить. Иначе говоря, на некоторых ионах решётки необходимо создать аномальный заряд. Это может быть осуществлено за счёт внутренних ресурсов решётки, путём пересаживания электронов с одних ионов решётки на другие. Так, например, в решётке

NaCl, построенной из ионов Na^+ и Cl^-, под действием освещения, электроны могут переводиться с отрицательных ионов хлора на положительные ионы натрия, в результате чего в решётке возникают нейтральные атомы натрия и нейтральные атомы хлора, являющиеся носителями, соответственно, электронной и дырочной проводимости.

Заметим, что наличие в решётке аномальных ионов является условием, необходимым, но недостаточным для возникновения электронной (дырочной) проводимости. Действительно, если аномальные ионы противоположных знаков находятся в непосредственном соседстве друг с другом, то они оказываются связанными между собой. Такая группа аномальных ионов, в целом электрически нейтральная, представляет собой образование, подобное френкелевскому экситону (этот термин впервые был введён Френкелем [14] в применении к гомеополярным решёткам). Экситон может перемещаться по решётке (это значит, что вся группа аномальных состояний, остающихся связанными между собой, перемещается, как одно целое), но не может дать тока. Чтобы сделать проводимость возможной, аномальные ионы необходимо развести настолько, чтобы можно было пренебречь взаимодействием между ними, т. е. чтобы можно было поведение каждого аномального иона рассматривать независимо от остальных.

С точки зрения зональной теории, создание в решётке двух свободных аномальных ионов с противоположными дополнительными зарядами означает перевод электрона из нормальной зоны в зону проводимости. Движение электрона в зоне проводимости или дырки в нормальной зоне свидетельствует о перемещении аномального состояния, которое последовательно передаётся с одного иона на соседний одноимённый ион. Если мы имеем дело с аномальным ионом, обладающим избыточным отрицательным зарядом, то мы говорим об электронном токе. Если же аномальный ион обладает избыточным положительным зарядом, то мы говорим о дырочном токе. В обоих случаях перемещение аномального состояния происходит за счёт перескакивания электрона с иона на ион. Таким образом, различие между электронной и дырочной проводимостью носит скорее терминологический, нежели физический характер.

В зависимости от того, какие ионы (металлические или металлоидные) служат передатчиками электрона или дырки, мы можем различать два механизма электронной и два механизма дырочной проводимости. В одних случаях электрону энергетически выгодно путешествовать по металлическим ионам: в других случаях — по металлоидным. То же самое можно сказать о дырке. Эти возможные механизмы схематически иллюстрированы на рис. 13. Они могут быть изображены также следующими символическими формулами.

Электронная проводимость:

$$1)\ M^{+p-1} + M^{+p} \longrightarrow M^{+p} + M^{+p-1},$$

$$2)\ R^{-q-1} + R^{-q} \longrightarrow R^{-q} + R^{-q-1}$$

Дырочная проводимость:

$$3)\ R^{-q+1} + R^{-q} \rightarrow R^{-q} + R^{-q+1},$$

$$4)\ M^{+p+1} + M^{+p} \rightarrow M^{+p} + M^{+p+1}.$$

Так, например, в решётке $CuO = Cu^{++} + O^{--}$ дырочный ток осуществляется аномальными ионами O^{--} (механизм № 3), в то время как в решётке $Cu_2O = Cu^{+} + O^{--}$ он осуществляется аномальными ионами Cu^{++} (механизм № 4).

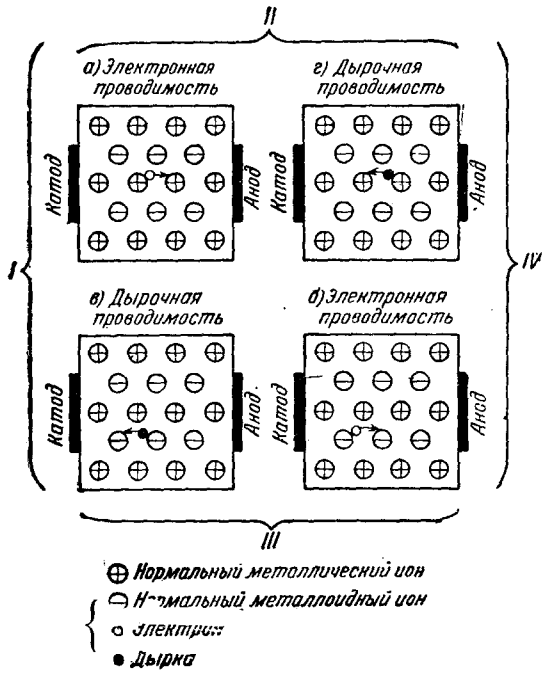

Рис. 13. Механизмы электронной проводимости.

Обратимся теперь к реальной гетерополярной решётке, обладающей механическими дефектами (междуузельные ионы или пустые узлы). Такая решётка уже с самого начала может содержать в себе некоторое количество аномальных ионов. Однако, в нормальном состоянии эти аномальные ионы локализованы около мест повреждений. Дефект вместе с прилипшим к нему аномальным ионом представляет собой электрически нейтральное образование, перемещение которого по решётке не создаёт тока. Таким образом, дефекты связывают аномальные ионы и этим исключают электронную (дырочную) проводимость. Аномальные ионы в свою очередь связывают дефекты, и тем самым исключается возможность ионной проводимости.

Электронная (дырочная) проводимость в реальной решётке может быть возбуждена не только в результате создания новых свободных аномальных ионов, но и в результате перевода уже существующих аномальных ионов из связанного состояния в свободное. Этот второй процесс требует меньших энергетических затрат, чем первый. На языке энергетической схемы рис. 10 такой перевод связанных

аномальных ионов в свободное состояние означает перевод электронов с электроположительных локальных уровней в зону проводимости или перевод дырок с электроотрицательных локальных уровней в нормальную зону.

§ 8. ВЛИЯНИЕ ПРИМЕСЕЙ И ТЕМПЕРАТУРЫ НА ЭЛЕКТРОПРОВОДНОСТЬ ПОЛУПРОВОДНИКОВ

По величине электропроводности полупроводники занимают промежуточное положение между металлами и изоляторами. Электропроводность металла имеет порядок величины 10^5 *ом*$^{-1}$ *см*$^{-1}$. Электропроводность изоляторов при обычных условиях порядка 10^{-15} *ом*$^{-1}$ *см*$^{-1}$. Полупроводники охватывают широкую промежуточную область. Электропроводность полупроводников колеблется в пределах от 10^2 до 10^{-10} *ом*$^{-1}$ *см*$^{-1}$. Таким образом, к группе полупроводников относятся тела с весьма разнообразными значениями электропроводности.

Электропроводность полупроводников существенно зависит от внешних условий. Воздействуя на полупроводник теми или иными физическими факторами, можно изменять его проводимость часто в миллионы раз. К таким факторам относятся:

а) примеси, б) температура, в) сильное электрическое поле, г) освещение.

Рассмотрим здесь роль примесей и температуры.

В полупроводниках, представляющих собой соединения типа $M_m R_r$, под наличием «примеси» обычно понимают всякое отклонение от стехиометрического соотношения. В этом случае примесным полупроводником называют всякий полупроводник, в котором стехиометрическое соотношение в той или иной степени нарушено, в отличие от чистого полупроводника, в котором стехиометрическое соотношение соблюдено.

В полупроводниках со стехиометрическим избытком металла проводимость имеет как правило электронную природу. В полупроводниках с избытком металлоида — дырочную природу. Заметим, что в некоторых случаях мы умеем делать полупроводник электронным или дырочным по нашему желанию, вызывая в нём стехиометрические отклонения в ту или иную сторону.

Для всех полупроводников характерна одна и та же закономерность: малейшие стехиометрические изменения вызывают резкие изменения в величине электропроводности. Электропроводность тем больше, чем сильнее нарушено стехиометрическое соотношение. Принято считать, что именно наличию «примесей» полупроводник обязан своею электропроводностью. «Чистые» полупроводники, с которыми, впрочем, мы никогда не имеем дела в действительности, с этой точки зрения являются по существу изоляторами. Эта мысль впервые была высказана Гудденом[15].

Резкая зависимость электропроводности от концентрации примеси может служить, повидимому, объяснением того, почему измерения различных авторов обычно столь резко расходятся между собой. Различные авторы при исследовании одного и того же материала, естественно, имеют дело с образцами, различным образом приготовленными и, следовательно, различающимися содержанием примеси. Электропроводность полупроводника не может считаться константой материала, подобно тому, как мы это считаем в отношении металлов и изоляторов.

Зависимость электропроводности от температуры для различных полупроводников различна. Различные типы этой зависимости схема-

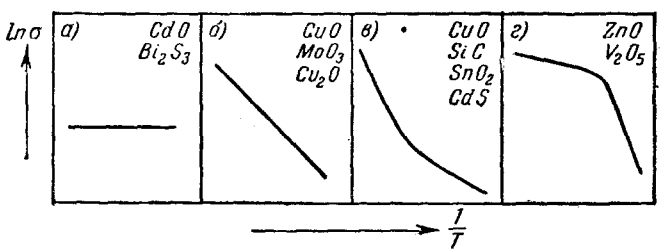

Рис. 14. Различные типы зависимости электропроводности от температуры.

тически изображены на рис. 14. На этом же рисунке выписаны некоторые характерные представители каждого типа. Таким образом, в неслишком широком температурном интервале электропроводность σ полупроводника всегда может быть выражена экспоненциальным законом:

$$\sigma = Ae^{-\frac{B}{T}}. \tag{8}$$

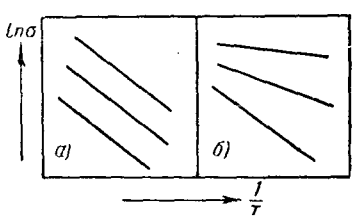

Рис. 15. Температурный ход электропроводности для образцов с различным содержанием примеси.

В частности, мы можем иметь здесь $B = 0$, что соответствует металлическому типу проводимости. Этот тип проводимости (рис. 14, *а*) встречается у полупроводников с достаточно большими σ.

Коэффициент A в (8) или вовсе не зависит от концентрации примеси или же возрастает вместе с возрастанием концентрации. В последнем случае это соответствует перемещению прямых на рис. 14 вверх параллельно самим себе, как это показано на рис. 15, *а*.

Коэффициент B также или вовсе не зависит от концентрации примеси, или же убывает при возрастании концентрации. В послед-

нем случае это соответствует уменьшению угла наклона прямых на рис. 14. По мере увеличения концентрации прямые становятся более пологими, как это изображено на рис. 15, *б*.

При переходе от одного температурного интервала к другому значения коэффициентов A и B меняются, что соответствует излому прямых на рис. 14. При переходе от низких температур в область высоких температур коэффициенты A и B у одних полупроводников увеличиваются (рис. 14, *в*), в то время как у других, наоборот, уменьшаются (рис. 14, *г*).

Различные типы температурной зависимости не связаны непосредственно с химической природой полупроводника. Различные образцы одного и того же полупроводника могут давать зависимости различного типа. Таким образом, для полупроводников не только величина электропроводности, но и температурный ход электропроводности не являются характеристикой материала, как это имеет место для металлов и изоляторов.

Теория электропроводности полупроводников, предложенная Вильсоном[16], представляет собой развитие идеи Гуддена, согласно которой электропроводность полупроводника целиком обусловливается дефектами кристаллической решётки. Теория относится к случаю, когда решётка содержит в себе дефекты лишь одного какого-либо сорта. Отсюда следует, что при соблюдённом стехиометрическом соотношении решётка должна считаться идеальной: дефекты в решётке свидетельствовали бы о нарушении стехиометрии.

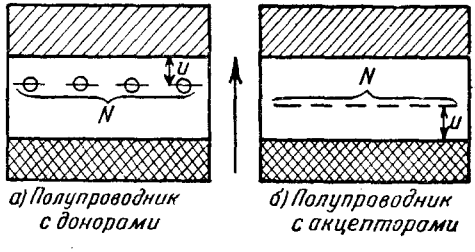

Рис. 16. Схема уровней для полупроводника с дефектами одного сорта (схема Вильсона).

Вильсон исходит из энергетической схемы, изображённой на рис. 16. В запрещённом участке между зонами мы имеем систему локальных уровней, которые являются или донорами (рис. 16, *а*) или акцепторами (рис. 16, *б*). Рис. 16, *а* относится к полупроводникам с избытком металла. Рис. 16, *б* — к полупроводникам с избытком металлоида. В первом случае температура перебрасывает электроны с локальных уровней в зону проводимости. Этими электронами, число которых экспоненциально возрастает с температурой, обусловливается электронная электропроводность полупроводника. Нормальная зона при этом вовсе не участвует в игре (это справедливо лишь при не слишком высоких температурах). Во втором случае температура перебрасывает электроны из нормальной зоны на локальные уровни. Дырки, возникающие при этом в нормальной зоне, обеспечивают дырочную проводимость

полупроводника. В этом случае исключается из игры зона проводимости (это справедливо, опять-таки, лишь при не слишком высоких температурах).

Теория Вильсона приводит к следующей формуле для электропроводности как функции температуры T и концентрации дефектов N:

$$\sigma = a\sqrt{Ne}^{-\frac{u}{2kT}} \tag{9}$$

Таким образом, теория Вильсона даёт правильный экспоненциальный закон (8) и объясняет зависимость константы A в (8) от концентрации дефектов. Параметр u в (9) имеет простой физический смысл: это есть энергия ионизации, т. е. энергия, затрачиваемая для перевода электрона (дырки) из связанного состояния в свободное (см. рис. 16, *а* и *б*).

Однако, теория Вильсона обладает следующими недостатками:

1) она не объясняет влияния концентрации дефектов на значение коэффициента B в экспоненциальном законе (8). В теории Вильсона коэффициент $B = \frac{u}{2k}$ целиком определяется природой решётки и природой дефектов;

2) характер зависимости коэффициента A от концентрации дефектов, к которому приводит теория Вильсона $A = a\sqrt{N}$, оправдывается на опыте далеко не всегда;

3) теория Вильсона не объясняет изломов на температурной кривой электропроводности;

4) полупроводники с металлическим ходом электропроводности остаются за рамками вильсоновской теории.

Некоторым обобщением теории Вильсона является теория Нейбура [17]. Нейбур имеет дело с полупроводником, энергетический спектр которого содержит одновременно как уровни-доноры, так и уровни-акцепторы, расположенные при этом на одной и той же горизонтали, как это изображено на рис. 17. Нижняя зона исключается из рассмотрения. У полупроводников такого типа, как это показал Нейбур, кривая электропроводности имеет вид, изображённый на рис. 14, *г*. При переходе из области высоких температур в область низких температур наклон кривой увеличивается вдвое. Такой излом на температурной кривой электропроводности с изменением наклона в два раза был обнаружен Яндером [18] в кристаллах ZnO. Формулы Нейбура имеют вид:

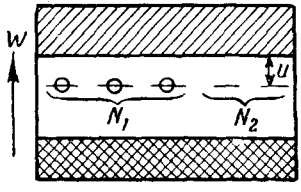

Рис. 17. Полупроводник с донорами и акцепторами (схема Нейбура).

при высоких температурах: $\sigma = a\sqrt{N_1}\, e^{-\frac{u}{2kT}}$,

при низких температурах: $\sigma = b\,\frac{N_1}{N_2}\, e^{-\frac{u}{kT}}$,

где

N_1 — концентрация доноров,

N_2 — концентрация акцепторов,

u — расстояние между локальными уровнями и зоной проводимости.

Теория Вильсона и теория Нейбура дают правдоподобные результаты для отдельных частных случаев, далеко не охватывая всей совокупности экспериментальных фактов.

Заметим, что, наряду с вильсоновской и нейбуровской схемами (рис. 16 и 17), можно было бы рассмотреть и более сложные энергетические схемы, характеризующиеся тем или иным размещением доноров и акцепторов в энергетическом спектре. Рассмотрение таких схем не встречает принципиальных затруднений. Однако, выбор той или иной схемы должен быть физически мотивирован.

Рассмотрим схемы, изображённые на рис. 10. Эти схемы превращаются в вильсоновскую схему в том частном случае, когда $N_1 = 0$ или $N_2 = 0$, и в нейбуровскую схему в том частном случае, когда $W_1 = W_2$. Полупроводники, изображающиеся схемами рис. 10, отличаются от вильсоновских полупроводников наличием в решётке двух сортов дефектов, а не одного (электроположительные и электроотрицательные дефекты). Число дефектов одного сорта непосредственно не зависит от числа дефектов другого сорта. Таким образом, решётка характеризуется не одним, а двумя независимыми между собой параметрами. Иначе говоря, степень нарушения стехиометрического соотношения и степень отклонения решётки от идеального состояния являются двумя самостоятельными характеристиками полупроводника. С точки зрения этой модели, в отличие от вильсоновской модели, восстановление в полупроводнике стехиометрического соотношения отнюдь ещё не означает возвращения решётки к идеальному состоянию.

Исходя из энергетических схем, изображённых на рис. 10, можно получить общую формулу для электропроводности σ, как функции температуры T и параметров N_1 и N_2 [19]. Эта формула имеет довольно сложный вид. Однако, на отдельных температурных участках, границы которых зависят от значений параметров N_1 и N_2, эта формула вырождается в простые экспоненциальные законы типа (8). Коэффициенты, фигурирующие в этом экспоненциальном законе, оставаясь постоянными внутри каждого такого участка, меняют свои значения при переходе от одного участка к другому, обусловливая тем самым изломы на температурной кривой. Варьируя значения параметров N_1 и N_2, можно получить семейство теоретических кривых. Такое семейство кривых схематически изображено на рис. 18. Оно охватывает все типы экспериментальных зависимостей σ от T.

Переход от одного прямолинейного участка на рис. 18 к другому прямолинейному участку связан с переменой механизма проводимости,

т. е. характеризуется включением в игру одних уровней и выключением других, так что подача электронов в зону проводимости (или дырок в нормальную зону) в разных областях температур происходит практически из различных резервуаров. Изгибы на кривых свидетельствуют отнюдь не о зависимости работы ионизации от температуры, а о наличии двух одновременно действующих источников ионизации (с определёнными, не зависящими от температуры, работами ионизации), относительная роль которых изменяется с изменением температуры. Металлический характер проводимости, наблюдающийся в некоторых случаях, получается из общей формулы как частный случай, соответствующий насыщению или, выражаясь точнее, истощению того резервуара, из которого черпаются электроны (или дырки) проводимости.

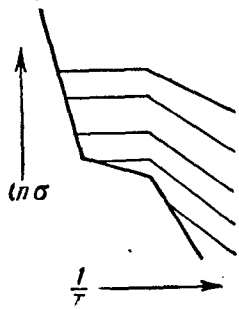

Рис. 18. Зависимость электропроводности от температуры (семейство теоретических кривых).

В теоретическую формулу для σ наряду с параметрами, характеризующими природу решётки, входят параметры N_1 и N_2, зависящие от биографии образца. Это значит, что электропроводность полупроводника определяется не только природой материала, но и технологией каждого данного образца. Следовательно, электропроводность полупроводника ни в коей мере не может рассматриваться как константа материала.

Все полупроводники, таким образом, могут быть описаны с помощью одной и той же энергетической модели. Качественные различия между ними имеют чисто количественное происхождение. Разнообразие экспериментальных кривых является результатом того, что различные экспериментаторы, имея дело с образцами, отличающимися друг от друга как своей природой, так и своей технологией, попадают на различные участки теоретической кривой.

§ 9. ДЕФОРМАЦИЯ ЭНЕРГЕТИЧЕСКИХ ЗОН ПОД ВЛИЯНИЕМ ВНЕШНЕГО ЭЛЕКТРИЧЕСКОГО ПОЛЯ

Если мы наложим на кристалл внешнее достаточно сильное электрическое поле, то поведение электрона в решётке делается весьма своеобразным.

На рис. 19, *а* ещё раз (ср. рис. 3, *а*) изображена кривая потенциальной энергии для электрона в решётке при отсутствии внешнего поля. При наложении поля кривая приобретает вид, изображённый на рис. 19, *б* (поле предполагается однородным и направленным по оси x). Мы видим, что поле вызывает два эффекта:

а) оно приподнимает потенциальный кратер каждого атома;

б) кроме того, оно возмущает каждый атом, т. е. искажает форму этого кратера.

Эти два эффекта имеют своим следствием деформацию энергетического спектра. Первый эффект приводит к наклонению энергетических зон. Второй эффект приводит к расширению зон.

Рассмотрим сперва эффект наклонения зон. Этот эффект (для одномерного случая) был исследован Зинером[20]. Наклонная зона изображена на рис. 20. На этом рисунке заштрихована та область значений W и x, внутри которой модуль волновой функции периодичен с периодом решётки. Эту область, в соответствии с терминологией, принятой в § 4, мы будем называть энергетической зоной. Вне зоны волновая функция быстро спадает, по мере удаления от границ зоны, так что вдали от зоны её можно считать практически равной нулю. Таким образом, мы получаем для электрона, строго говоря, сплошной энергетический спектр. Все значения энергии возможны. Каждому значению энергии соответствует определённая волновая функция, которая, однако, практически отлична от нуля лишь в ограниченной области пространства. Схема, изображённая на рис. 20, отнюдь не выражает собой зависимости энергии W от координаты x. Она должна восприниматься всего лишь как характеристика волновой функции.

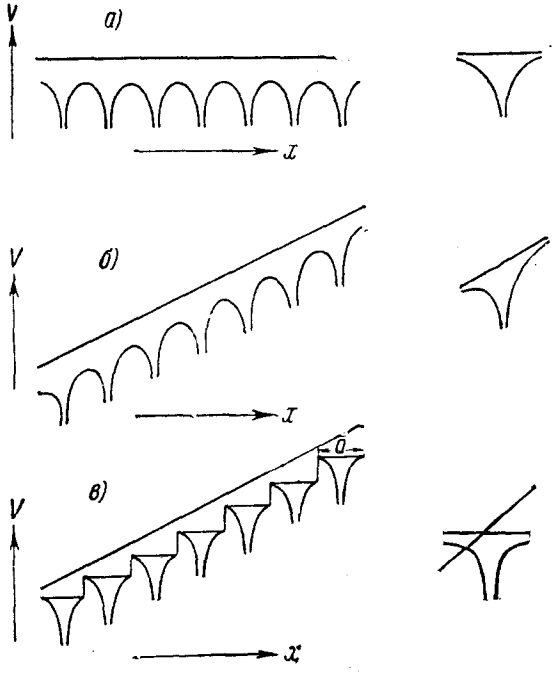

Рис. 19. Кривая потенциальной энергии для электрона в решётке.

а) при отсутствии поля, *б*) при наличии поля (истинная кривая), *в*) при наличии поля (апроксимированная кривая).

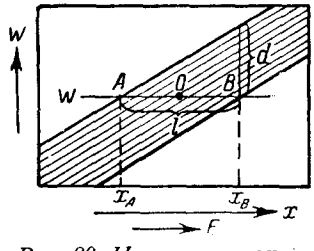

Рис. 20. Наклонная зона.

Электрон с заданной энергией W оказывается запертым внутри зоны. Иначе говоря, он локализован в более или менее узкой области пространства, ограниченной плоскостями $x = x_A$ и $x = x_B$ (см. рис. 20). Эта область тем уже, чем сильнее поле.

Действительно, если d есть ширина наклонной зоны в «вертикальном» направлении, а l — ширина зоны в «горизонтальном» направлении, то

$$l = \frac{d}{eE},$$

где e — абсолютная величина заряда электрона, E — напряжённость электрического поля.

Таким образом, как это ни странно на первый взгляд, по мере усиления поля электрон делается в известном смысле всё более и более связанным. Такой электрон совершает колебания между левой и правой границами зоны (точки A и B на рис. 20). В правой (нижней) половине зоны (отрезок OB на рис. 20) электрон обладает положительной эффективной массой и, следовательно, ускоряется в сторону действующей на него силы, т. е. справа налево на рис. 20. В левой (верхней) половине зоны (отрезок OA на рис. 20) эффективная масса электрона оказывается отрицательной. Попадая в эту область, электрон обнаруживает свойства, парадоксальные с классической точки зрения: он начинает ускоряться в сторону, прямо противоположную действующей на него силе, т. е. слева направо на рис. 20. Таким образом, если при отсутствии поля электрон был связан с определённым уровнем внутри зоны, т. е. мог быть охарактеризован тройкой квантовых чисел n_1, n_2, n_3, то теперь электрон начинает путешествовать по уровням зоны. Его состояние уже не может быть охарактеризовано какой-либо определённой тройкой значений n_1, n_2 n_3.

Обратимся теперь к эффекту расширения зон[21]. Этот эффект обусловлен штарковским расщеплением соответствующих термов изолированного атома. Всякий вырожденный терм можно рассматривать как совокупность нескольких термов, при отсутствии поля совпадающих друг с другом. При наложении поля эти совпадающие термы раздвигаются, причём сдвиг каждого из них, вообще говоря, различен. В результате мы получаем расщепление терма, как это показано на рис. 21. Энергетическую зону, соответствующую такому вырожденному терму, можно представлять себе как результат наложения нескольких зон, число которых равно кратности вырождения. Расщепление терма вызывает раздвижение соответствующих зон, в результате чего зона в целом оказывается расширенной. Этот эффект поясняется рис. 22, на котором изображены две соседние зоны, из которых нижняя соответствует невырожденному s-терму, а верхняя — p-терму, обладающему трёхкратным вырождением.

Если внешнее поле достаточно слабо, так что потенциал поля мало меняется при переходе от одного атома к соседнему, то потенциал внутри каждого атома можно считать постоянным. В этом случае потенциальная кривая, изображённая на рис. 19, *б*, может быть апроксимирована ступенчатой кривой, изображённой на рис. 19, *в*.

При такой апроксимации эффект расширения зон оказывается потерянным.

Расширение зон имеет своим следствием сужение запрещённого участка между зонами, как это видно из рис. 22. Таким образом, учёт штаркэффекта приводит к сближению зон, возрастающему по мере увеличения поля.

Это обстоятельство должно отражаться прежде всего на оптическом спектре кристалла, приводя к деформации полос поглощения. Так, например, полоса, соответствующая переходам электронов из

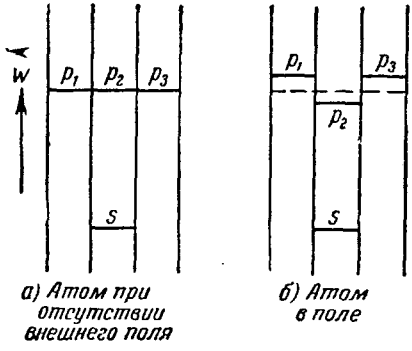

Рис. 21. Штарковское расщепление термов в изолированном атоме.

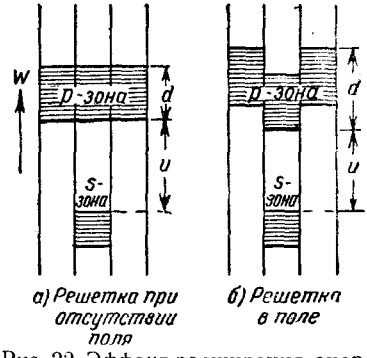

Рис. 22. Эффект расширения энергетических зон.

нормальной зоны в зону проводимости (так называемая собственная полоса), при наложении поля должна расширяться, причём красная граница этой полосы, как это может показаться на первый взгляд, должна сдвигаться в красную сторону спектра. В действительности, однако, сдвиг красной границы в красную сторону отнюдь не всегда обязан иметь место и не может рассматриваться как критерий штарковского расширения зон. В этом можно убедиться, если учесть те правила отбора, которыми регулируются оптические перебросы электронов из одной зоны в другую.

Согласно этим правилам, с каждого данного уровня данной зоны возможен переход лишь на один вполне определённый уровень другой зоны, так что, если все уровни внутри каждой зоны считать перенумерованными, то оптический переход может осуществляться лишь между двумя уровнями с одинаковыми номерами. Существенно при этом, что нумерация уровней должна проводиться для одних зон снизу вверх, а для других — сверху вниз, в зависимости от природы зоны. На рис. 23, *а* представлены две соседние зоны (нормальная зона и зона приводимости) с одинаковым порядком нумераций уровней, а на рис. 23, *б* — с противоположным порядком нумерации. Стрелка № 1 изображает переход, соответствующий красной границе

полосы поглощения; стрелка № 2 соответствует фиолетовой границе. В случае, изображённом на рис. 23, *а*, красная граница определяется переходом с нижнего уровня нижней зоны на нижний уровень верхней зоны. Этот случай имеет место, например, при сочетании *s*-зоны с *p*-зоной. В случае, изображённом на рис. 23, *б*, красная граница обусловливается переходом с верхнего уровня нижней зоны на нижний уровень верхней зоны. Этот случай имеет место, например, при комбинации *s*-зоны с *d*-зоной.

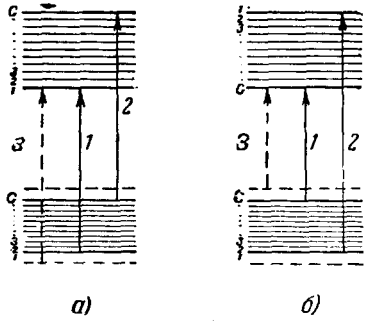

Рис. 23. Правила отбора при оптических переходах между зонами.
↑ *1* — красная граница полосы поглощения при отсутствии поля, ↑ *2* — фиолетовая граница при отсутствии поля, ↑ *3* — красная граница при наличии поля.

Предположим, что при наложении внешнего поля расширение испытывает лишь нижняя зона (границы такой расширенной зоны представлены на рис. 23 пунктирными линиями). Это имеет место тогда, когда верхняя зона является *s*-зоной. Пунктирная стрелка на рис. 23, *а* и 23, *б* изображает красную границу полосы поглощения при наличии поля. Мы видим, что при включении поля красная граница оказывается сдвинутой в красную сторону в случае, изображённом на рис. 23, *б*, в то время как в случае, изображённом на рис. 23, *а*, она, наоборот, оказывается сдвинутой в фиолетовую сторону, хотя в обоих случаях запрещённый участок между зонами при включении поля оказывается суженным.

§ 10. ИОНИЗИРУЮЩЕЕ ДЕЙСТВИЕ ЭЛЕКТРИЧЕСКОГО ПОЛЯ

Чтобы вызвать в решётке проводимость, необходимо воздействовать на неё такими агентами, которые снабжали бы электронами зону проводимости или опустошали бы нормальную зону. Такими ионизирующими агентами являются температура и освещение. Внешнее электрическое поле, будучи достаточно сильным, также может выступать в роли самостоятельного ионизирующего агента. Ионизирующая способность поля делается возможной благодаря той деформации, которую оно вызывает в энергетическом спектре решётки. Эта деформация, как мы видели, сводится к наклонению и расширению зон.

Благодаря наклонению зон, электрон приобретает способность переходить с одного уровня на другой через разделяющий их запрещённый участок двумя путями, которые отмечены на рис. 24 вертикальной и горизонтальной стрелками *3* и *4*. Вертикальный переход связан с затратой энергии; этот переход аналогичен возбуждению.

Переходы этого рода обусловлены механизмом ударной ионизации (ионизация столкновениями). Горизонтальный переход происходит без затраты энергии; он подобен просачиванию через потенциальный барьер. Такие переходы мы будем называть электростатической ионизацией. Поле, вызывая наклонение зон, делает возможными переходы как того, так и другого рода. В обоих случаях ионизирующая способность поля тем больше, чем у́же запрещённый участок, который должен быть преодолён электроном.

Как горизонтальные, так и вертикальные переходы могут совершаться:

1) с локальных уровней на зону проводимости (чисто электронный ток);

2) с нормальной зоны на локальные уровни (чисто дырочный ток);

3) с нормальной зоны на зону проводимости (смешанный ток).

Горизонтальные переходы, кроме того, могут совершаться:

4) с катода на зону проводимости (чисто электронный ток);

5) с нормальной зоны на анод (чисто дырочный ток).

Таким образом, ионизация может развиваться или во всём объёме кристалла (случаи 1, 2, 3) или только в его приэлектродных слоях (случаи 4, 5).

Благодаря расширению зон, запрещённый участок между зонами сужается (как в горизонтальном, так и в вертикальном направлении). На рис. 24 границы такой расширенной зоны изображены пунктирными линиями. Таким образом, штарковское расширение служит фактором, облегчающим ионизацию при любом из механизмов. Учитывая штаркэффект, мы тем самым заставляем оба ионизационных механизма (ударная ионизация, электростатическая ионизация) вступать в действие при полях, несколько более слабых, чем это имеет место при пренебрежении эффектом Штарка. Таким образом, сужение запрещённого участка оказывает такое же действие, как дополнительное увеличение поля. Кроме того, штаркэффект усиливает роль температурной ионизации (что может быть существенным в случае полупроводников). Действительно, сужение запрещённого участка увеличивает концентрацию тепловых электронов в зоне проводимости или дырок в нормальной зоне. Здесь сужение запрещённого участка ока-

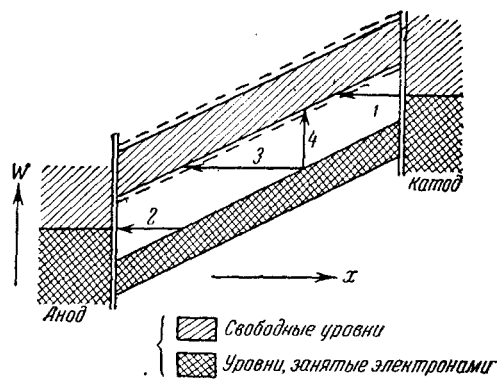

Рис. 24. Ионизирующие действия поля.

зывает такое же действие, как дополнительное повышение температуры.

Рассмотрим подробнее механизм ударной ионизации. Вертикальные переходы происходят в результате соударений между электронами, так что энергия одного электрона увеличивается за счёт другого. При этих процессах зона проводимости обогащается электронами или нормальная зона обогащается дырками, или же одновременно происходит и то и другое.

В первом случае свободный электрон (ионизирующий электрон), двигаясь в зоне проводимости, сталкивается с электроном, сидящим на локальном уровне (ионизуемый электрон) и подымает его в зону проводимости, теряя при этом как раз ту энергию, которую приобретает электрон ионизуемый. Ионизовавший электрон продолжает оставаться в зоне проводимости, падая, однако, с более высоких её уровней на более низкие. Очевидно, такого рода соударения свободный электрон может совершать лишь тогда, когда, двигаясь по зоне проводимости, он уходит достаточно далеко от её нижней границы.

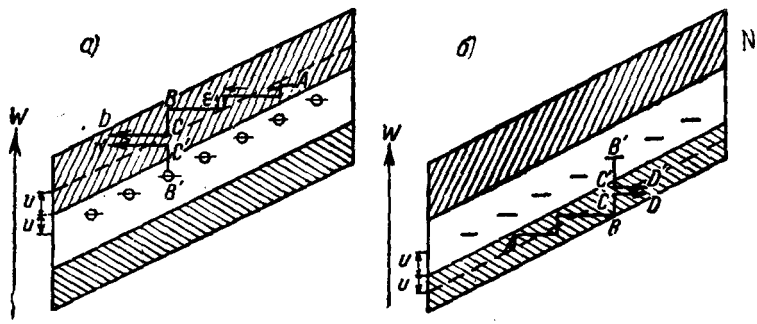

Рис. 25. Механизм ударной ионизации.

а) ионизация ударом электрона, *б*) ионизация ударом дырки, $ABCD$ — путь ионизирующего электрона (дырки), скачок BC соответствует моменту ионизации. $B'C'D'$—путь ионизуемого электрона (дырки). Очевидно: $B'C' = BC$.

Расстояние его от нижней границы должно быть не меньше u (если считать по вертикали, см. рис. 25, *а*), где u — ширина запрещённого участка между нижней границей зоны проводимости и тем уровнем, на котором сидит ионизуемый электрон. Это имеет место лишь при достаточно сильных полях. Действительно, свободный электрон, двигаясь к аноду, спускается по ступенькам лестницы, лежащей внутри зоны проводимости (см. рис. 25, *а*); каждый переход с одной ступеньки на следующую нижнюю свидетельствует об акте соударения с решёткой. Длина ступеньки λ есть средняя длина свободного пробега, а высота ступеньки ε есть та средняя порция энергии, которую теряет электрон при каждом таком соударении. Будет ли заходить этот ступенчатый путь электрона в ту область зоны проводимости, где делается возможной ударная ионизация, или будет целиком ле-

жать ниже этой области — это зависит (при заданных λ и ε) от того, насколько сильно наклонена зона, т. е. от величины внешнего поля E. В рассматриваемом случае, когда ионизация происходит при помощи свободного электрона, двигающегося по зоне проводимости, энергия ионизирующего электрона до соударения, конечно, больше энергии ионизуемого. После соударения энергия ионизовавшего электрона может оказаться как больше, так и меньше, чем энергия электрона ионизованного.

Во втором случае ионизация происходит при помощи дырки, двигающейся в нормальной зоне. В действительности же здесь сталкиваются два электрона, оба принадлежащие нормальной зоне, в результате чего один из них (ионизуемый) выбрасывается вверх за пределы нормальной зоны на пустой локальный уровень, а другой (ионизирующий), оставаясь в нормальной зоне, падает на её более низкий уровень, который предполагается свободным. Очевидно, что такое соударение может происходить лишь при условии, что дырка находится достаточно далеко от верхней границы нормальной зоны. Действительно, двигаясь к катоду, дырка поднимается по ступенькам лестницы, лежащей внутри нормальной зоны (рис. 25, *б*). Будет ли эта лестница заходить в требуемую для ионизации область — это зависит от того, насколько сильно наклонена зона. В рассматриваемом случае до соударения между электронами энергия ионизирующего электрона не обязана быть больше энергии ионизуемого. После же соударения энергия ионизовавшего электрона (т. е. оставшегося в нормальной зоне), конечно, заведомо меньше энергии ионизованного (т. е. выброшенного из зоны).

Наконец, в третьем случае мы имеем столкновения между свободным электроном зоны проводимости и связанным электроном нормальной зоны. При таком столкновении связанный электрон делается свободным и, таким образом, зона проводимости, обогащаясь электронами, тем самым обогащает дырками нормальную зону.

Наряду с процессом ударной ионизации, при котором электрон, так или иначе, совершает вертикальный переход снизу вверх, имеет место обратный процесс, при котором электрон совершает такой же вертикальный переход, но сверху вниз, отдавая свою энергию решётке. Этот процесс приводит к связыванию свободного электрона (дырки). Если бы этого процесса не существовало, то ток в кристалле возрастал бы с расстоянием подобно лавине. Стационарный ток, постоянный по всему объёму кристалла, устанавливается лишь тогда, когда оба процесса (ионизация столкновением и связывание) приходят в равновесие. Учёт процесса связывания представляет собой как раз то, чем современные теории ударной ионизации [22, 23] отличаются от старых теорий.

Обратимся теперь к механизму электростатической ионизации. Горизонтальные переходы на рис. 24, т. е. переходы без изменения энергии, становятся возможными в результате наклонения зон. Дей-

ствительно, волновая функция для электрона с заданной энергией быстро затухает в запрещённом участке; однако, её хвост простирается достаточно далеко. Это и обусловливает то обстоятельство, что электрон всегда имеет некоторый шанс очутиться по ту сторону запрещённого участка.

Вероятность такого просачивания чрезвычайно чувствительна к длине запрещённого участка, преодолеваемого электроном. Этот участок укорачивается по мере наклонения зон, т. е. по мере возрастания поля. Вероятность горизонтального перехода, будучи практически равной нулю при не слишком сильных полях, почти внезапно становится весьма значительной, коль скоро зоны приобретают достаточный наклон. Таким образом, при достаточно сильном поле зона проводимости всегда может ожидать, что в ней появится электрон, просочившийся в неё с локального уровня или непосредственно с катода или, наконец, из нормальной зоны. Нормальная зона в свою очередь всегда может ожидать, что в ней возникнет дырка за счёт электрона, ушедшего сквозь тоннель на пустой локальный уровень или непосредственно на анод или, наконец, в зону проводимости.

Таким образом, кристалл может обогащаться свободными электронами и дырками, с одной стороны, за счёт своих собственных ресурсов, а с другой стороны — за счёт металлических электродов (стрелки *1* и *2* на рис. 24). Роль, которую играют в кристалле его собственные электроны и дырки, рассмотрена в работе Зинера[20]. Роль, которую могут играть чужие электроны и дырки, сорванные с поверхности электродов, рассматривается в работе Фаулера[25]. Теория Фаулера имеет преимущество перед теорией Зинера, так как она требует преодоления более короткого барьера.

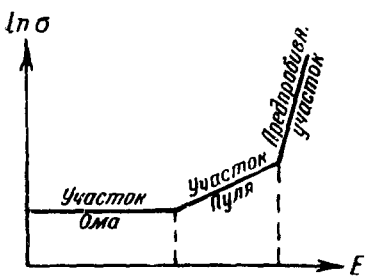

Рис. 26. Электропроводность как функция поля.

Кривая электропроводности как функция поля изображена на рис. 26. Это — экспериментальная кривая. Она распадается на три участка, сменяющих друг друга по мере возрастания поля:

а) участок закона Ома,

б) участок закона Пуля,

в) предпробивной участок.

Экспоненциальное возрастание проводимости на участке Пуля не может быть объяснено с помощью рассмотренных механизмов ионизации. Ударная ионизация приводит к неправильной температурной зависимости. Электростатическая ионизация даёт неправильную зависимость от E (более быстрое возрастание проводимости с полем, чем это имеет место в действительности). Что касается предпробивного участка, то в этой области электростатическая ионизация может рассматриваться как один из возможных механизмов.

В механизме ударной ионизации, так же как и в механизме электростатической ионизации, электрическое поле выступает в качестве самостоятельного ионизирующего агента. Оба эти механизма становятся возможными, как мы видели, благодаря наклонению зон. Рассмотрим ещё один механизм ионизации, в котором ионизирующим агентом является температура, а поле выступает в качестве помощника. Здесь вспомогательная роль поля делается возможной благодаря штарковскому расширению зон.

Для проводимости, обусловленной температурной ионизацией, мы имеем:

$$\sigma = Ce^{-\frac{U}{2kT}}, \tag{9}$$

где U — ширина запрещённого участка между уровнями, поставляющими электроны (дырки), и уровнями, принимающими на себя электроны (дырки). При учёте штаркэффекта величина U должна рассматриваться как функция поля:

$$U(E) = U(0) - \Delta U(E). \tag{10}$$

Зависимость ΔU от E может быть получена из теории штаркэффекта[26]. Она выражается следующей эмпирической формулой:

$$\Delta U(E) = aE^2 + bE. \tag{11}$$

В этой формуле параметры a и b сами медленно меняются с полем, причём при возрастании E отношение $\frac{a}{b}$ убывает, так что в слабых полях можно пренебречь вторым членом в (11) по сравнению с первым (область квадратичного штаркэффекта), а в сильных полях, наоборот, можно пренебречь первым членом по сравнению со вторым (область линейного штаркэффекта).

Подставляя (10) в (9) и учитывая (11), мы получим в области сильных полей:

$$\sigma = Ae^{\alpha E}. \tag{12}$$

Эта формула выражает собой закон Пуля. Коэффициент α имеет правильный порядок величины. Кроме того, зависимость α от температуры T также имеет правильный характер. Таким образом, эффект Штарка является, повидимому, одним из возможных механизмов, обусловливающих возрастание проводимости на участке Пуля[27].

В заключение заметим, что облегчение термической ионизации под влиянием поля может происходить не только в результате штаркэффекта, но и за счёт совершенно иного механизма, рассмотренного в работе Френкеля[28]. Механизм Френкеля подобен механизму Шоттки, вызывающему понижение работы выхода при термической эмиссии электронов из металла. С точки зрения теории Френкеля, работа U, необходимая для того, чтобы пересадить электрон с одного иона решётки на другой и развести образовавшиеся таким образом ано-

мальные ионы на бесконечность, должна в присутствии внешнего электрического поля E уменьшаться на величину

$$\Delta U = 2\sqrt{\frac{e^3 E}{\varepsilon}},$$

где ε — диэлектрическая постоянная кристалла. Таким образом, теория Френкеля приводит к выражению

$$\sigma = Be^{\beta\sqrt{E}} \tag{13}$$

вместо выражения (12), основанного на механизме штаркэффекта. Формула Френкеля (13) отличается от эмпирического закона Пуля заменой E на $\sqrt{E}$. Это обстоятельство, однако, не следует расценивать как возражение против теории, ибо нельзя сказать, чтобы имеющиеся экспериментальные данные свидетельствовали бы против такой замены.

§ 11. ФОТОЭЛЕКТРИЧЕСКИЕ ПРОЦЕССЫ

Спектр поглощения реального кристалла содержит в себе как правило две системы полос:

а) так называемые собственные полосы поглощения, лежащие в коротковолновой части спектра;

б) так называемые примесные полосы поглощения в длинноволновой части спектра.

Полосы этих двух типов обладают существенно различными свойствами. Собственные полосы довольно устойчивы: они лишь в очень слабой степени реагируют на внешние воздействия. Примесные же полосы представляют собой весьма хрупкие образования: их структура чрезвычайно чувствительна к внешним воздействиям. В некоторых образцах эти полосы могут вовсе отсутствовать. Они могут быть созданы искусственно и искусственно уничтожены. Под действием некоторых факторов они могут превращаться друг в друга, т. е. поглощение кристалла может возрастать в одной примесной полосе за счёт ослабления поглощения в другой примесной полосе. Коэффициент поглощения в примесных полосах в среднем значительно меньше, чем в собственных полосах.

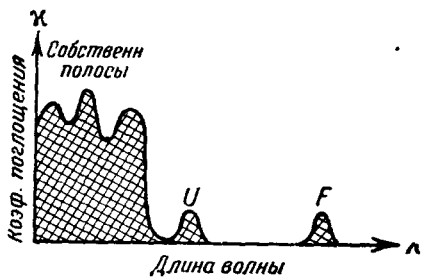

Рис. 27. Спектр поглощения щёлочно-галоидных кристаллов.

Рассмотрим некоторые свойства полос поглощения у щёлочно-галоидных кристаллов. Спектр поглощения для этих кристаллов имеет вид, изображённый на схематическом рис. 27. По оси ординат отложен коэффициент поглощения; по оси абсцисс — длины волн.

Собственные полосы лежат в далёком ультрафиолете. Мы имеем несколько собственных полос, наложенных друг на друга. Они образуют, таким образом, сплошной спектр с отдельными максимумами, следующими друг за другом.

Что касается примесных полос, то в щёлочно-галоидных кристаллах можно наблюдать две основные примесные полосы. Одна из них лежит в ультрафиолетовой части спектра, вблизи собственных полос. Это — так называемая *U-полоса*. Другая — в видимой части спектра. Эту полосу обычно называют *F-полосой*. Этой полосой определяется окраска кристалла. При отсутствии *F*-полосы кристалл в видимой части спектра прозрачен.

U- и *F*-полосы отличаются друг от друга прежде всего своими фотоэлектрическими свойствами. *F*-полоса обладает значительным фотоэффектом: поглощение света в этой полосе приводит к возникновению проводимости. *U*-полоса, наоборот, фотоэлектрически нечувствительна. Положение этих полос в спектре у разных щёлочно-галоидных кристаллов не зависит от химической природы кристалла, а всецело определяется постоянной решётки. Мы имеем следующие эмпирические формулы:

$$\text{для } F\text{-полосы } \nu a^2 = \text{const},$$
$$\text{для } U\text{-полосы } \nu a = \text{const},$$

где a — постоянная решётки, ν — частота, соответствующая максимуму полосы.

Примесные полосы могут превращаться друг в друга. Если кристалл, содержащий *U*-полосу, но не содержащий *F*-полосы, освещать частотами, принадлежащими *U*-полосе, то *U*-полоса постепенно будет исчезать, а за счёт неё будет возникать *F*-полоса. Кристалл будет постепенно окрашиваться. Обратный процесс происходит самопроизвольно. Он протекает тем быстрее, чем выше температура.

F-полоса может быть получена также другими способами. Существует три таких способа.

1. Освещение кристалла рентгеновскими лучами или ультрафиолетовым светом из длинноволновой части собственной полосы. Это — так называемый субстрактивный способ окрашивания. При длительном освещении достигается насыщение: *F*-полоса достигает своей максимальной в данных условиях интенсивности, не меняющейся при дальнейшем освещении. Эта максимальная интенсивность тем больше, чем больше дефектов содержится в решётке и чем выше температура, при которой происходит окрашивание. Кристаллы, окрашенные таким субстрактивным способом, самопроизвольно обесцвечиваются. Это обесцвечивание может быть ускорено повышением температуры или наложением внешнего электрического поля.

2. Прогревание кристалла в парах щёлочного металла. Это — так называемый аддитивный способ окрашивания. Образующаяся при этом *F*-полоса определяется природой кристалла, но не природой

того щёлочного металла, в парах которого происходит прогревание. В кристалле, окрашенном таким способом, при наложении поля вся окрашенная область движется к аноду, оставляя за собой кристалл бесцветным.

3. Наконец, можно окрасить кристалл с помощью электрического поля. При острых электродах и достаточно высокой температуре окраска появляется в области катода. Она как бы выходит из катода и постепенно распространяется в глубь кристалла. При этом одновременно с F-полосой возникает также и U-полоса. По мере продвижения окрашенного блока в направлении анода его интенсивность ослабевает; в то же время интенсивность U-полосы возрастает. При перемене направления поля движение окрашенной области изменяется на обратное. F-полоса таким путём может быть загнана обратно в электрод. Кристалл вновь оказывается бесцветным. Однако, U-полоса при этом застревает в кристалле.

При субстрактивном окрашивании окрашивающие центры в кристалле создаются за счёт его собственных ресурсов. При аддитивном окрашивании они вводятся извне. В этом смысле создание F-полосы в результате освещения U-частотами может рассматриваться как один из видов субстрактивного окрашивания. Окрашивание электрическим полем является одним из видов аддитивного окрашивания.

Все эти своеобразные свойства полос поглощения щёлочно-галоидных кристаллов в настоящее время очень хорошо изучены. Мы имеем несколько сот экспериментальных работ. Сюда относятся работы школы Поля[29] в Германии, а также целая серия работ П. С. Тартаковского[30] и его учеников в СССР.

Что касается полупроводников, то их спектр поглощения обладает той же структурой, что и спектр поглощения щёлочно-галоидных кристаллов. Основное различие состоит в том, что при переходе к полупроводникам все полосы сдвигаются в красную сторону спектра. Собственные полосы попадают в видимую часть. Кристаллы полупроводников, таким образом, непрозрачны. Примесные полосы оказываются соответственно перемещёнными в красную и инфракрасную области, мало удобные для экспериментального исследования. Именно это обстоятельство является, повидимому, причиной того, почему спектры полупроводников до сих пор изучены очень плохо. Мы имеем здесь лишь несколько случайных работ.

Спектры поглощения реальных гетерополярных кристаллов в общих чертах могут быть интерпретированы с точки зрения энергетических схем, представленных на рис. 10. Всякий поглощённый квант перебрасывает электрон с одного уровня на другой вышележащий уровень. Примесные полосы, с точки зрения этих схем, соответствуют переходам электронов с локальных уровней в зону проводимости, или из нормальной зоны на локальные уровни. Собственные полосы соответствуют переходам из зоны в зону. Такая интерпретация приводит

к следующим следствиям, находящимся в качественном согласии с экспериментальными данными:

1. Интенсивность примесных полос с этой точки зрения должна определяться не только природой кристалла, но и степенью отклонения нашей реальной решётки от идеального состояния и степенью отклонения от стехиометрического соотношения. Иначе говоря, интенсивность этих полос должна находиться в непосредственной зависимости от обработки образца, и, таким образом, может быть различной у различных образцов. Интенсивность собственной полосы, напротив того, определяется исключительно природой кристалла, чем и обусловливается её устойчивость. Эти свойства полос поглощения находятся, как мы видели, в соответствии с действительностью.

2. Примесные полосы должны лежать в более длинноволновой части спектра, нежели собственные полосы. Это заключение вытекает непосредственно из рассмотрения рис. 10. Кроме того, примесные полосы должны быть значительно менее интенсивными, чем собственные полосы. Действительно, интенсивность примесных полос при нашей интерпретации обусловлена числом дефектов в решётке, в то время как интенсивность собственных полос — числом регулярных атомов. Эти свойства полос также соответствуют действительности.

3. Освещение кристалла частотами, принадлежащими какой-либо примесной полосе, может привести в некоторых случаях, как это видно из анализа рис. 10, к перекачиванию электронов с одних локальных уровней на другие локальные уровни (например, через зону проводимости). Действительно, электроны, срываемые светом с одних уровней, попав в зону проводимости и странствуя в ней, могут прилипать к другим уровням. В результате — доноры становятся акцепторами, а акцепторы превращаются в доноров. Интенсивность нашей примесной полосы ослабляется за счёт возникновения другой примесной полосы. Кристалл переводится в своего рода метастабильное состояние, которое постепенно будет рассасываться. Такого рода процессы, как мы видели, наблюдаются в действительности.

Заметим, однако, что интерпретация полос поглощения с точки зрения схем, изображённых на рис. 10, на первый взгляд встречает на своём пути следующее затруднение. С точки зрения этих схем поглощение света в любой полосе приводит к образованию электронов в зоне проводимости или дырок в нормальной зоне или к тому и другому одновременно. Таким образом, все полосы поглощения должны быть фотоэлектрически активными, т. е. поглощение света во всех случаях должно сопровождаться внутренним фотоэффектом. В действительности это не так. Одни полосы поглощения обладают значительным фотоэффектом, в то время как в других полосах фотоэффект практически совсем отсутствует. Это имеет место, как мы уже отмечали, в щёлочно-галоидных кристаллах, в которых, при сильном фотоэффекте в F-полосе, поглощение света в U-полосе не вызывает сколько-нибудь заметной проводимости. Такое фотоэлектри-

ческое неактивное поглощение наблюдается также и в некоторых полупроводниках (например в V_2O_5).

Это обстоятельство послужило поводом для развития Я. И. Френкелем представлений об экситоне[14]. Действительно, такое фотоэлектрически неактивное поглощение естественно сопоставить с непроводящими возбуждёнными состояниями решётки, когда два аномальных иона, возникших в результате поглощения кванта, остаются связанными между собой и, таким образом, не могут участвовать в переносе тока.

Представление об экситонах не является, однако, обязательным для объяснения фотоэлектрически неактивного поглощения. Этот эффект может быть интерпретирован в рамках обычных энергетических схем, с которыми мы оперируем в теории полупроводников. Здесь существенно лишь, чтобы в энергетическом спектре кристалла в качестве локальных уровней фигурировали бы одновременно как уровни-доноры, так и уровни-акцепторы (см. рис. 10, 17). Можно показать[31], что свет может переводить электроны в зону проводимости и, тем не менее, при определённых условиях этот процесс может и не вызывать сколько-нибудь заметного увеличения проводимости. Такой перевод электронов в свободное состояние, не сопровождающийся увеличением проводимости, обусловлен изменениями в условиях рекомбинации, происходящими под действием света.

Заметим в заключение, что, наряду с нормальным фотоэлектрическим эффектом и фотоэлектрически неактивным поглощением света, в некоторых полупроводниках при определённых условиях имеет место аномальный (так называемый отрицательный) фотоэффект, когда под влиянием освещения проводимость не увеличивается, а уменьшается. Такой отрицательный фотоэффект наблюдался в селене, стибните, молибдените и закиси меди. Результаты исследования весьма разноречивы. Общим, однако, является тот факт, что отрицательный фотоэффект, если вообще он имеет место, наблюдается лишь при не слишком низких температурах и в достаточно сильных электрических полях. При понижении температуры или при понижении поля (при неизменных прочих условиях) отрицательный эффект исчезает и обращается в нормальный эффект. Критическая температура, при которой фотоэффект меняет знак, лежит тем ниже, чем больше напряжённость поля.

На основании имеющихся экспериментальных данных пока трудно решить, обусловлен ли отрицательный фотоэффект каким-либо самостоятельным механизмом или же является результатом вторичных процессов, накладывающихся на нормальный эффект. Заметим, однако, что сам факт уменьшения проводимости под влиянием освещения может быть объяснён с точки зрения энергетических схем, изображённых на рис. 10[32]. Здесь существенно, чтобы световые и тепловые электроны были бы различны по своему происхождению, т. е. чтобы свет черпал электроны с иных уровней, чем тепло. В этом

случае свободные электроны получают возможность рекомбинировать с дырками двух сортов: с дырками, созданными светом, и с дырками, созданными теплом. Соответствующие коэффициенты рекомбинации не обязаны быть одинаковыми. При определённом соотношении между этими коэффициентами, условия рекомбинации при включении освещения могут оказаться значительно облегчёнными, в результате чего суммарная электропроводность, при наличии освещения, может оказаться меньше, чем в темноте.

ЛИТЕРАТУРА

1. Волькенштейн. Электропроводность полупроводников, ГТТИ, 1946.
2. Bloch. *Zs. f. Phys.*, **52**, 555, 1929.
3. Sommerfeld. *Zs. f. Phys.* **47**, 1, 1928; **47**, 43, 1928.
4. Drude. *Ann. d. Phys.*, **1**, 566, 1900; **3**, 370, 1900; **7**, 687, 1902.
5. Jones and Zener. *Proc. Roy. Soc.* A, **144**, 101, 1934.
6. Wilson. *Proc. Roy. Soc.* A, **133**, 458, 1934.
7. Блохинцев. *ЖЭТФ*, **5**, 213, 1935.
8. Brillouin. La méthode du champ self-consistent, Paris, 1933.
9. Tamm. *Sow. Phys.*, **1**, 733, 1932.
10. Frenkel. *Sow. Phys.*, **9**, 158, 1936.
11. Landau. *Sow. Phys.* **3**, 664, 1933.
12. Schottky. *Zs. f. Phys. Chem.*, **29**, 335, 1935.
13. Frenkel. *Sow. Phys.*, **12**, 1, 1936.
14. Frenkel. *Sow. Phys.*, **8**, 185, 1935.
15. Gudden. *Erlangen Ber.*, **62**, 289, 1930.
16. Wilson. *Proc. Roy. Soc.*, A, **134**, 277, 1932.
17. Nijboer. *Proc. Phys. Soc.*, **51**, 575, 1939.
18. Jander. *Zs. Anorgan. u. Allgemein. Chem.*, **192**, 286, 1930; **199**, 165, 1931.
19. Волькенштейн. *ЖЭТФ*, 1946 (печатается).
20. Zener. *Proc. Roy. Soc.* A, **145**, 523, 1934.
21. Волькенштейн. *Труды ФИАН СССР*, вып. 2, 1936.
22. Писаренко. *Изв. АН. СССР*, 631, 1938.
23. Fröhlich. *Proc. Roy. Soc.*, A. **160**, 230, 1937.
24. Волькенштейн. *ЖТФ*, **5**, 583, 1935.
25. Fowler. *Proc. Roy. Soc.*, A, **141**, 56, 1933.
26. Stark. *Hb. d. Exp. Phys.*, т. 21, 468—502.
27. Волькенштейн. *ЖТФ*, **9**, 171, 1939.
28. Френкель. *ЖТФ*, **8**, 1292, 1938.
29. Gudden. Lichtelektrische Erscheinungen. Berlin 1928.
30. Тартаковский. Внутренний фотоэффект в диэлектриках, ГТТИ, 1940.
31. Волькенштейн. *ЖЭТФ*, **10**, 1229, 1940.
32. Волькенштейн. *ЖЭТФ*, 1946 (печатается).