Подчеркнем, что речь идет в данном случае только о внешних полях. Движение электронов в периодическом поле идеальной решетки, безусловно, должно рассматриваться квантовомеханически. Все нужные нам характеристики этого движения, однако, сводятся к закону дисперсии $E_l(\mathbf{p})$ (в частном случае (1.10″) — к эффективной массе m). Коль скоро закон дисперсии задан, дальнейшие рассуждения, касающиеся поведения системы во внешних полях, могут быть классическими. При этом уравнения (1.10) — (1.10″) дают практически всю необходимую информацию.

§ 2. Электроны и дырки

Уравнения (1.10) — (1.10″) в сочетании с формулой для тока (1.4) позволяют придать ясный смысл представлению о «положительных дырках», введенному в гл. I на основании опытных данных о знаке постоянной Холла и термоэдс. Для этой цели заметим прежде всего, что, как мы знаем, эффективные массы, входящие в (1.10) — (1.10″), могут быть и отрицательными: так обстоит дело вблизи потолка энергетической зоны. Легко видеть, что это можно интерпретировать как изменение знака заряда носителя. Действительно, пусть $\mathbf{F}$ есть сила Лоренца

$$\mathbf{F} = e\mathscr{E} + \frac{e}{c}[\mathbf{v}\times\mathfrak{B}].$$

Видно, что в уравнения движения (1.10) — (1.10″) входят не заряд электрона и компоненты тензора $m_{\alpha\beta}^{-1}$ по отдельности, а только определенная комбинация их: $em_{\alpha\beta}^{-1}$ или, в случае (1.10″), e/m. Отсюда явствует, что электрон с отрицательной эффективной массой ускоряется электрическим и магнитным полями как частица с положительной массой и положительным зарядом.

Отрицательные значения эффективной массы не должны вызывать удивления. Действительно, уравнения (1.10′), (1.10″) описывают среднее ускорение электрона под действием силы Лоренца. Кроме нее, однако, есть еще сила, действующая на электрон со стороны атомов решетки и учитываемая видом закона дисперсии $E_l(\mathbf{p})$. Именно она и ответственна за возникновение отрицательной эффективной массы.

Направление макроскопических потоков заряда, энергии и т. д. в системе электронов зависит от того, какие именно электроны дают главный вклад в них. Здесь следует выделить предельные случаи почти свободной и почти заполненной зоны.

Первый из них соответствует зоне проводимости в полупроводнике обычного типа. В условиях, близких к термодинамическому равновесию, все электроны находятся у дна зоны, где эффективные массы положительны. Только эти электроны и могут участвовать в явлениях переноса — мы имеем здесь систему отрицательно заряженных носителей.

Второй случай отвечает, например, валентной зоне, из которой удалено некоторое (небольшое) число электронов (они могут быть, в частности, переброшены в зону проводимости за счет энергии теплового движения решетки). В условиях, близких к термодинамически равновесным, не заняты только состояния у потолка зоны. Нижние, полностью заполненные состояния, не дают вклада в потоки заряда, энергии и т. д. Действительно, в силу (III.4.4) каждому электрону с квазиимпульсом $\mathbf{p}$ там можно поставить в соответствие другой электрон с квазиимпульсом $-\mathbf{p}$ и той же энергией (III.4.4). Согласно (1.4) вклады этих электронов в ток взаимно компенсируются. Отличный от нуля вклад в поток электронов могут дать только те из них, которые не имеют «партнеров» с той же энергией и противоположно направленным квазиимпульсом. Это — электроны, находящиеся в не полностью занятых состояниях, т. е. у потолка зоны, где эффективные массы отрицательны. Иначе говоря, мы имеем здесь как бы систему положительно заряженных носителей.

При этом весьма удобным оказывается представление о «дырках» в не полностью заполненной зоне. Его легко ввести, замечая, что, как было выше показано, плотность тока $\mathbf{j}_l$, создаваемая всеми электронами целиком заполненной зоны, равна нулю:

$$\mathbf{j}_l = 2\sum_{\mathbf{p}} \mathbf{j}_l(\mathbf{p}) = 0. \tag{2.1}$$

Здесь множитель 2 возник из-за двух возможных значений проекции спина (при заданном квазиимпульсе $\mathbf{p}$); суммирование по $\mathbf{p}$ охватывает первую зону Бриллюэна. Пусть теперь один электрон удален из зоны, в результате чего освободилось состояние с квазиимпульсом $\mathbf{p}'$ и, скажем, положительной проекцией спина *) (состояние с квазиимпульсом $\mathbf{p}'$ и отрицательной проекцией спина при этом остается занятым). Тогда плотность тока будет

$$\mathbf{j}_l = 2\sum_{\mathbf{p}\neq\mathbf{p}'} \mathbf{j}_l(\mathbf{p}) + \mathbf{j}_l(\mathbf{p}') = 2\sum_{\mathbf{p}} \mathbf{j}_l(\mathbf{p}) - \mathbf{j}_l(\mathbf{p}') = -\mathbf{j}_l(\mathbf{p}'). \tag{2.2}$$

Плотность тока, создаваемая электронами зоны с одним незанятым состоянием, равна по величине и противоположна по направлению плотности тока, которая создавалась бы недостающим электроном.

Согласно (1.4) и (I.3.13) при $\mathfrak{B}=0$ плотность тока, связанная с одним электроном в единице объема, есть

$$\mathbf{j}_l(\mathbf{p}') = \frac{e^2}{m}\langle\tau\rangle\,\mathfrak{E}.$$

*) Направление, на которое проектируется спин, здесь может быть выбрано произвольно.

В рассматриваемом случае $m < 0$. Следовательно,

$$\mathbf{j}_l = -j_l(\mathbf{p}') = \frac{e^2}{|m|}\langle\tau\rangle\mathcal{E}.$$

Это та плотность тока, которая создавалась бы частицей с массой $|m|$ и зарядом e, движущимся со скоростью $\frac{e}{|m|}\langle\tau\rangle\mathcal{E}$.

Таким образом, дело обстоит так, как если бы носителями заряда были частицы с положительной массой $|m|$ и положительным зарядом e. Фактически перемещаются, разумеется, электроны. Однако электрон у потолка зоны движется против действующей на него силы. По этой причине удаление такого электрона приводит к увеличению плотности тока — такому, как если бы появилась «частица» с положительной массой $|m|$ и положительным зарядом. Иначе говоря, коллективное движение всех оставшихся электронов валентной зоны эквивалентно движению одной «частицы».

Аналогично обстоит дело и с плотностью потока энергии. Она дается выражением

$$\mathbf{I}_l = \sum_{\mathbf{p},\,\sigma} E_l(\mathbf{p})\,\mathbf{v}(\mathbf{p})\,n(\mathbf{p},\ \sigma). \tag{2.3}$$

Здесь $n(\mathbf{p},\ \sigma)$ — концентрация электронов с квазиимпульсом $\mathbf{p}$ и проекцией спина σ. В полностью заполненной зоне $n(\mathbf{p},\ \sigma) = 1/V$, где V — объем системы.

При одном недостающем электроне в единичном объеме, плотность потока энергии будет, подобно (2.2),

$$\mathbf{I}_l = 2\sum_{\mathbf{p}\neq\mathbf{p}'} E_l(\mathbf{p})\,\mathbf{v}_l(\mathbf{p}) + E_l(\mathbf{p}')\,\mathbf{v}_l(\mathbf{p}') = -E_l(\mathbf{p}')\,\mathbf{v}_l(\mathbf{p}'). \tag{2.4}$$

Поток энергии, создаваемый электронами зоны с одним незанятым состоянием, равен по величине и противоположен по направлению потоку энергии, который создавался бы недостающим электроном. Иначе говоря, он равен потоку энергии, создаваемому одной «частицей» с энергией $-E_l(\mathbf{p}')$.

Совершенно так же можно было бы рассмотреть и случай нескольких отсутствующих электронов (с квазиимпульсами $\mathbf{p}'$, $\mathbf{p}''$, ...). Потоки заряда и энергии здесь оказываются такими же, как если бы они переносились положительно заряженными «частицами» (с энергиями $-E_l(\mathbf{p}')$, $-E_l(\mathbf{p}'')$, ...), взятыми в таком числе, сколько электронов отсутствует.

Эти «частицы» получили название «дырок», ибо происхождение их связано с наличием незаполненных состояний («мест») в зоне *). При описании явлений переноса «электронный» и «дырочный» языки полностью эквивалентны. Последний язык особенно удобен, если речь идет о почти заполненной зоне. Здесь, как мы видели,

*) Представление о дырках было введено Я. И. Френкелем (1928 г.)

свободны состояния у потолка зоны E_v и энергии дырок, отсчитанные от него, равны $-E_l$ (p). В частности, при изотропном параболическом законе дисперсии мы имеем

$$-E_l(\mathbf{p}) = -E_v - \frac{p^2}{2m} = -E_v + \frac{p^2}{2|m|}.$$

Как и следовало ожидать, это есть энергия частицы с положительной эффективной массой $|m|$. Энергия дырки у потолка валентной зоны равна $-E_v$. Эффективные массы дырок обычно снабжают индексом p (не путать с квазиимпульсом), а электронов — n. В данном случае $m_p = |m_n|$.

В такой ситуации, интересуясь только явлениями переноса заряда и энергии, мы можем вообще «забыть» о существовании валентной зоны и электронов в ней. Взамен вводится представление о газе дырок и о дырочной зоне. Минимум энергии в последней совпадает с потолком валентной зоны, а энергия дырок на зонной схеме отсчитывается вниз (рис. 4.1). Квазиимпульс дырки равен квазиимпульсу отсутствующего электрона с обратным знаком. Иначе говоря, мы вправе рассматривать газ носителей заряда двух знаков с положительными эффективными массами.

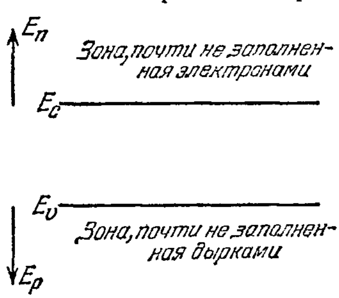

Рис. 4.1. Направления отсчета энергий электронов проводимости E_n (p) и дырок E_p (p).

В связи с этим результатом следует обратить внимание на три обстоятельства.

Во-первых, аналогия между дырками и электронами не полна. Действительно, электроны могут существовать и в вакууме, в то время как представление о дырках имеет смысл лишь постольку, поскольку существуют валентные оболочки, в которых могут оказаться незаполненные места. В отсутствие коллектива электронов представление о дырках лишается смысла. По этой причине дырки часто называют квазичастицами.

Во-вторых, эквивалентность описания системы с помощью представлений об электронах или дырках доказана выше только для потоков заряда и энергии, или, общее, с точки зрения поведения системы во внешних полях. Отсюда, однако, не следует, что эти картины эквивалентны и с точки зрения поля, создаваемого самими этими электронами и дырками. Иначе говоря, «электронная» и «дырочная» картины эквивалентны кинематически, но не всегда — динамически. Вопрос о том, какая из двух картин «правильнее», будет обсужден позднее, в гл. XVII.

Наконец, в-третьих, предыдущие рассуждения относились к поведению системы в электрическом и магнитном полях, но отнюдь не в гравитационном (и, следовательно, инерционном) поле. Понятие об эффективной массе, а с ним и понятие о положительных дырках не имеют никакого отношения к гравитации (и инерции). Суть дела состоит в том, что силы инерции определяются через производную по времени от импульса, а не от квазиимпульса. Поэтому, например, в известных опытах Толмэна и Стюарта (определение удельного заряда электронов по измерению импульса тока при торможении вращающейся проволочной катушки) отношение заряда частиц к массе всегда будет получаться равным $-e/m_0$, т. е. будет соответствовать отрицательным электронам с истинной (а не эффективной) массой m_0. Такой результат получится и в том случае, когда вместо металла будет взят дырочный полупроводник. Далее, в § III.5 мы видели, что эффективная масса зависит от степени перекрытия атомных волновых функций. Она сильно увеличивается по мере увеличения глубины энергетической зоны; для глубоких узких зон $m \gg m_0$ (внутренние электроны атомов практически локализованы около своих ядер). Однако это обстоятельство никак не сказывается на весе тела, который определяется истинными массами всех ядер и электронов.

§ 3. Движение носителей заряда в постоянном и однородном магнитном поле (классическая теория). Диамагнитный резонанс

Рассмотрим задачу о движении электрона или дырки в постоянном и однородном магнитном поле. Решение этой задачи приведет нас к важному методу экспериментального определения эффективной массы. Дабы не затемнять существа дела, пренебрежем сначала рассеянием носителей заряда. Позднее мы учтем и этот эффект и выясним, в чем состоит его роль и в какой мере она существенна в рассматриваемой задаче.

Направим оси координат вдоль главных осей тензора обратной эффективной массы. Тогда, принимая во внимание формулу для силы Лоренца

$$\mathbf{F} = -\frac{e}{c}[\mathbf{v} \times \mathcal{B}],$$

мы можем, применительно к рассматриваемой задаче, переписать уравнения (1.10′) в виде

$$\begin{aligned} \frac{dv_x}{dt} &= -\frac{\mathcal{B}e}{cm_x}(\alpha_z v_y - \alpha_y v_z), \\ \frac{dv_y}{dt} &= -\frac{\mathcal{B}e}{cm_y}(\alpha_x v_z - \alpha_z v_x), \\ \frac{dv_z}{dt} &= -\frac{\mathcal{B}e}{cm_z}(\alpha_y v_x - \alpha_x v_y). \end{aligned} \qquad (3.1)$$