Журнал экспериментальной и теоретической физики

Т. 43 *1962* *Вып. 3(9)*

ТЕОРИЯ ЭЛЕКТРОННОГО ЭНЕРГЕТИЧЕСКОГО СПЕКТРА МЕТАЛЛОВ С РЕШЕТКОЙ ТИПА ВИСМУТА

А. А. Абрикосов, Л. А. Фальковский

На основе нескольких качественных предположений и теории симметрии определяется электронный энергетический спектр металлов пятой группы периодической системы.

1. Как известно, металлы пятой группы периодической системы, т. е. мышьяк, сурьма и висмут, обладают рядом интересных особенностей. Кристаллические решетки этих металлов однотипны. Они относятся к ромбоэдрической системе с двумя атомами в ячейке. Особенностью этих решеток является то, что они могут быть получены из простых кубических решеток путем незначительных смещений атомов. Это продемонстрировано на рис. 1.

Возьмем простую кубическую решетку и выделим две гранецентрированные подрешетки. Вершина куба одной из них находится в центре куба другой. Атомы разных подрешеток обозначены черными и светлыми кружками. Теперь сместим подрешетки относительно друг друга вдоль пространственной диагонали куба. Симметрия решетки (класс) при этом сразу станет ромбоэдрической. В дополнение к этому надо произвести небольшую ромбоэдрическую деформацию самих подрешеток. Возникающая решетка и есть решетка типа висмута. В таблице приведены числа, которые характеризуют относительное расположение подрешеток (ромбоэдрические координаты двух атомов, относящихся к разным подрешеткам, равны (s, s, s), $(-s, -s, -s)$) и угол между основными периодами (для куба речь идет о подрешетке). Из этих данных видно, насколько решетки мышьяка, сурьмы и висмута близки к простой кубической.

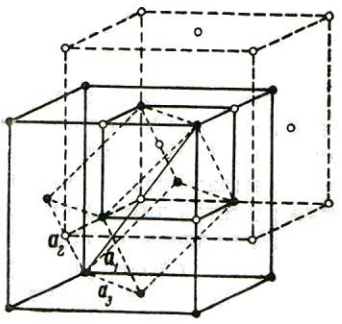


Рис. 1

Другая особенность этих элементов заключается в том, что они являются плохими металлами, т. е. число носителей тока, приходящиеся на один атом, в них чрезвычайно мало. Так у висмута, согласно данным по эффекту де Гааза — ван Альфена [1, 2], число «свободных» электронов составляет $\sim 10^{-5}$ на один атом. Согласно Люттингеру [3], полная плотность электронов выражается формулой $n = 2(2\pi)^{-3}(pV_0 + V_F)$, где V_0 — объем ячейки обратной решетки, V_F — объем, ограниченный всеми ферми-поверхностями (со стороны меньших энергий), отнесенной к одной ячейке, p — целое число. Эта формула имеет тот же самый вид, что и для модели невзаимодействующих частиц в периодическом поле. Так как $V_0 = (2\pi)^3/v_0$, где v_0 — объем элементарной ячейки, то из формулы Люттингера следует, что если число электронов, приходящихся на одну элементарную ячейку, нечетно, то вещество обязательно будет металлом, если же оно четно, то вещество может быть как металлом, так и диэлектриком. В последнем случае число носителей типа

	s	θ		s	θ
As	0,226	54°10′	Bi	0,237	54°14′
Sb	0,223	57°6′	Куб	0,25	60°

«электронов» должно равняться числу «дырок». Отсюда следует, что у рассматриваемых металлов обязательно должны быть носители обоих типов.

Помимо определения числа носителей тока, из данных по эффекту де Гааза — ван Альфена можно получить довольно детальные сведения об электронном спектре. По-видимому, поверхность Ферми для электронов имеет форму трех эллипсоидов (или близких к ним фигур) с центрами на бинарных осях [1]. В последнее время на основе опытных данных об инфракрасном поглощении [4] высказано мнение об отклонении формы этих поверхностей от точного эллипсоида [5]. Что же касается дырок, то их поверхность Ферми, по-видимому, вообще не похожа на эллипсоид [2].

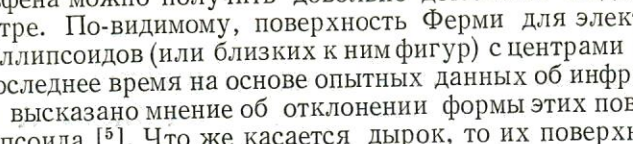


Рис. 2

Настоящая работа посвящена теории электронного спектра металлов пятой группы. Она основывается на форме решетки этих металлов и некоторых качественных предположениях. Получающиеся при этом результаты дают возможность привести аргументы в пользу закономерности возникновения решеток типа висмута.

Отметим здесь, что попытка связать форму решетки висмута с малым числом носителей тока уже предпринималась Джонсом [6]. Однако она не идет дальше качественных рассуждений.

2. Одним из авторов [7] было показано, что для нахождения спектра возбуждений в металле можно действовать в полной аналогии с моделью невзаимодействующих частиц в периодическом поле. Ввиду этого мы будем основываться на такой аналогии. В рассматриваемых нами металлах число электронов является нечетным. Следовательно, если бы они имели простую кубическую решетку, они были бы хорошими металлами. В результате деформации период решетки меняется, число электронов в ячейке удваивается и металлы в принципе могут стать диэлектриками.

Такой переход проще всего понять на примере одномерной модели металла. Предположим, что путем незначительного изменения потенциала мы увеличили его период вдвое. Это немедленно приведет к расщеплению энергетических полос. Из рис. 2 видно, что максимум $\varepsilon(\mathbf{k})$ в нижней из получающихся полос и минимум в верхней возникнут при совершенно определенных значениях квазиимпульса, которые характеризуются следующими свойствами (мы рассматриваем одну ячейку исходной обратной решетки): а) в исходной обратной решетке это точки, которым соответствуют одинаковые значения энергии; б) в новой обратной решетке это эквивалентные точки, т. е. отличающиеся на период обратной решетки, в то время как в старой решетке они не были эквивалентными. Даже если в исходной решетке число электронов на атом было нечетным, то новое вещество будет диэлектриком.

В трехмерном случае такое категорическое утверждение является несправедливым, и в новой решетке минимумы верхней полосы могут оказаться ниже максимумов верхней. Более того, если деформация бесконечно мала, то вещество обязательно останется металлом. Это следует из того, что очень

малое расщепление не сможет изменить площади ферми-поверхности (хотя, конечно, изменит число носителей). Однако в том случае, если расщепление является конечным, вещество может стать диэлектриком. Мы не можем привести исчерпывающего доказательства того, что это обязательно должно иметь место при той ситуации, которую мы рассматриваем. Однако можно высказать некоторые соображения.

Прежде всего отметим, что одно лишь изменение периода может значительно уменьшить число носителей. Например, если предположить, что в исходной решетке ферми-поверхность близка к сфере с объемом, равным половине объема ячейки обратной решетки, то появляющиеся при изменении периода новые грани зоны Бриллюэна в пространстве обратной решетки отсекают от нее объемы порядка 0,01. Можно предположить, что увеличение деформации приведет к уменьшению этих объемов. С другой стороны, металл с малым числом носителей может перейти в диэлектрик [1]. Здесь очевидно существенно соотношение среднего расстояния между квазичастицами и радиуса экранировки кулоновского взаимодействия. Если радиус экранировки велик по сравнению со средним расстоянием между квазичастицами, то неэкранированное кулоновское взаимодействие может привести к образованию связанных пар из возбуждений электронного и дырочного типа. Очевидно, при этом металл переходит в диэлектрик. Нетрудно увидеть, что критерий перехода в диэлектрик может быть записан в виде $v \ll 1/ma$, где v — скорость на ферми-границе, m — масса свободного электрона, a — период решетки. При спектре типа $\varepsilon = p^2/2m^*$, $v \sim (N/V)^{1/3} m^{*-1}$, откуда следует, что неустойчивость возникает при малом числе носителей.

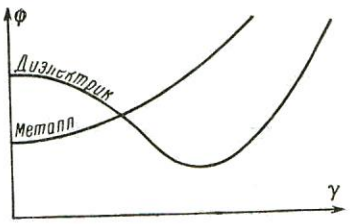


Рис. 3

Переход из металла в диэлектрик является фазовым переходом первого рода [2]. Если рассматривать термодинамические потенциалы обеих фаз как функции деформации γ, то соответствующие кривые должны иметь вид, изображенный на рис. 3. Если минимум термодинамического потенциала диэлектрической фазы ниже энергии металлической фазы при $\gamma = 0$, то такая деформация наступит самопроизвольно.

Для нас пока существенно то, что выше и ниже точки пересечения кривых на рис. 3 должны существовать обе фазы, причем одна из них будет равновесной, а другая метастабильной. Если предположить, что в случае висмута имеет место именно эта ситуация, то мы можем заключить, что «диэлектрическая» фаза существует при сколь угодно малой деформации в качестве метастабильной фазы. В точке, где деформация обращается в нуль, «диэлектрическая» фаза становится абсолютно неустойчивой.

Название «диэлектрическая» мы берем в кавычки. Это объясняется тем, что в действительности эта фаза в решетке типа висмута, как мы увидим ниже, будет не диэлектриком, а плохим металлом с малым числом электронов. При этом оказывается, что скорость на ферми-границе имеет тот же порядок величины, что и в хороших металлах, благодаря чему существование таких малых групп не противоречит изложенным соображениям.

Для получения энергетического спектра мы сначала рассмотрим решетку, бесконечно близкую к кубу, а затем будем деформировать ее и переведем в истинную решетку висмута. При этом для того, чтобы изменение спектра

[1] Изложенные ниже соображения основаны на представлениях, высказанных Р. Г. Архиповым.

[2] На первый взгляд может показаться, что неравновесность металлической фазы является только следствием уменьшения числа носителей и может возникнуть даже при бесконечно малой деформации. В действительности это не так, ввиду того что скорость на поверхности Ферми при бесконечно малой деформации имеет прежний порядок величины. Непрерывный переход при конечной деформации исключается, ввиду того что образование пар приводит к уменьшению экранировки и поэтому, начавшись, должно идти до конца.

было непрерывным, без фазового перехода, мы все время будем рассматривать «диэлектрическую» фазу, хотя она и является неравновесной при малых деформациях.

Если в исходной кубической решетке выделить две гранецентрированные подрешетки, то их периоды будут равны

$$\mathbf{a}_1 = \mathbf{a}_2^0 + \mathbf{a}_3^0, \quad \mathbf{a}_2 = \mathbf{a}_3^0 + \mathbf{a}_1^0, \quad \mathbf{a}_3 = \mathbf{a}_1^0 + \mathbf{a}_2^0. \tag{1}$$

Периоды новой обратной решетки будут равны

$$\mathbf{K}_1 = 2\pi\,[\mathbf{a}_2\mathbf{a}_3] / (\mathbf{a}_1\,[\mathbf{a}_2\mathbf{a}_3]) = \tfrac{1}{2}(\mathbf{K}_2^0 + \mathbf{K}_3^0 - \mathbf{K}_1^0),$$

$$\mathbf{K}_2 = \tfrac{1}{2}(\mathbf{K}_3^0 + \mathbf{K}_1^0 - \mathbf{K}_2^0), \quad \mathbf{K}_3 = \tfrac{1}{2}(\mathbf{K}_1^0 + \mathbf{K}_2^0 - \mathbf{K}_3^0). \tag{2}$$

Произвольный вектор обратной решетки равен

$$\mathbf{K} = \tfrac{1}{2}(n_3 + n_2 - n_1)\,\mathbf{K}_1^0 + \tfrac{1}{2}(n_3 + n_1 - n_2)\,\mathbf{K}_2^0 + \tfrac{1}{2}(n_1 + n_2 - n_3)\,\mathbf{K}_3^0. \tag{3}$$

Рассмотрим точки, лежащие на пространственных диагоналях исходного куба посредине между его вершиной и центром. Таких точек всего восемь, причем все они неэквивалентны в старой обратной решетке. В новой решетке они разбиваются на четыре пары, в каждую из которых входят по две эквивалентные (т. е. отличающиеся на новый период обратной решетки) точки ($\mathbf{k}_i$, $-\mathbf{k}_i$). Координаты точек $\mathbf{k}_i$ можно записать посредством коэффициентов в формуле $\mathbf{k} = \sum_s \alpha_s \mathbf{K}_s^0$:

$$\mathbf{k}_0 = \left(\tfrac{1}{4}\ \tfrac{1}{4}\ \tfrac{1}{4}\right), \quad \mathbf{k}_1 = \left(\tfrac{1}{4}\ \tfrac{\bar{1}}{4}\ \tfrac{\bar{1}}{4}\right), \quad \mathbf{k}_2 = \left(\tfrac{\bar{1}}{4}\ \tfrac{1}{4}\ \tfrac{\bar{1}}{4}\right),$$

$$\mathbf{k}_3 = \left(\tfrac{\bar{1}}{4}\ \tfrac{\bar{1}}{4}\ \tfrac{1}{4}\right). \tag{4}$$

Особые свойства точек $\mathbf{k}_i$, $-\mathbf{k}_i$ — неэквивалентность в исходной решетке и эквивалентность в новой, а также высокая симметрия — дают основания ожидать появления в них максимумов и минимумов энергии как функции квазиимпульса (см., например, рис. 2 для одномерной модели).

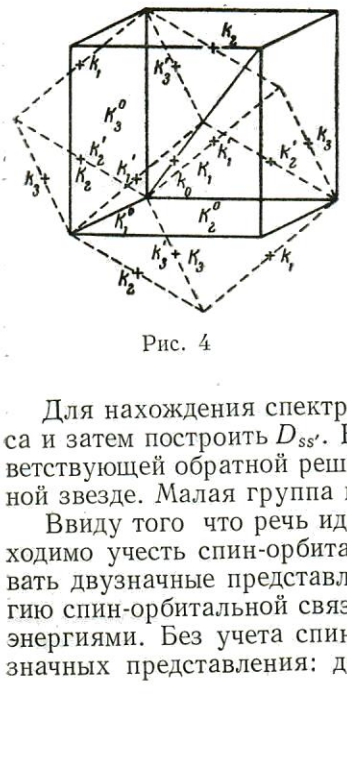


Рис. 4

Расположение точек $\mathbf{k}_i$ в обратной решетке проиллюстрировано на рис. 4. После деформации точка $\mathbf{k}_0$ оказывается в центре ромбоэдра. Малая группа этой точки есть D_{3d} и она сама по себе образует звезду. Точки $\mathbf{k}_1$, $\mathbf{k}_2$ и $\mathbf{k}_3$ принадлежат к одной звезде. Они расположены на бинарных осях ромбоэдра. Малая группа $\mathbf{k}_1$ есть C_{2h}. Отметим, что расположение этих точек в точности соответствует экспериментально обнаруженному расположению центров ферми-поверхностей.

Для нахождения спектра надо, прежде всего, определить функции базиса и затем построить $D_{ss'}$. Начнем с исходной кубической решетки. В соответствующей обратной решетке все восемь точек $\mathbf{k}_i$, $-\mathbf{k}_i$ принадлежат к одной звезде. Малая группа каждой точки есть C_{3v}.

Ввиду того что речь идет о металлах с большим атомным номером, необходимо учесть спин-орбитальную связь, а следовательно, надо рассматривать двузначные представления группы C_{3v}. Мы будем предполагать энергию спин-орбитальной связи малой по сравнению с обычными электронными энергиями. Без учета спин-орбитальной связи группа C_{3v} имеет три однозначных представления: два одномерных и одно двумерное (см. [8]). При

учете спина соответствующие представления должны быть умножены на $D_{1/2} = E_2'$ (см. [9]). При этом оба одномерных представления переходят в эквивалентные двумерные, двузначные представления E_2', а двумерное дает одно двумерное представление E_2' и два комплексно сопряженных одномерных представления E_1', объединяемых в одно.

Ввиду того что решетка содержит центр инверсии, двукратное вырождение сохраняется при всех значениях $\mathbf{k}$. Поэтому для тех представлений, которые получены из одномерных, учет спин-орбитальной связи не приводит к расщеплению уровней. Иное положение имеет место для двумерного однозначного представления, где спин-орбитальная связь расщепляет четырехкратный уровень на два двукратных.

Если предположить, что рассматриваемый уровень соответствовал одномерному (однозначному) представлению группы C_{3v}, то можно показать (с помощью метода, аналогичного примененному ниже), что после деформации никак не может возникнуть спектр того типа, какой наблюдается у висмута. Ввиду этого будем исходить из двумерного представления.

Теперь, однако, необходимо учесть, что мы рассматриваем решетку с измененным периодом. В новой решетке точки $\mathbf{k}_i$ и $-\mathbf{k}_i$ становятся эквивалентными. При этом соответствующие представления должны быть объединены. В результате матрица $D_{ss'}$ в окрестности каждой точки $\mathbf{k} = \mathbf{k}_i$ становится матрицей порядка 8×8.

Для оценки порядка величин коэффициентов матрицы удобно ввести ψ_+ и ψ_- — функции базиса двумерного (однозначного) представления группы C_{3v}, обладающие свойствами $C_3\psi_\pm = e^{\pm 2\pi i/3}\psi_\pm$, $\sigma_{(1)}\psi_\pm = \psi_\mp$ ($\sigma_{(1)}$—отражение в плоскости, проходящей через ось C_3) и спиноры $\chi_\pm$, преобразующиеся по правилу $C_3\chi_\pm = e^{\pm i\pi/3}\chi_\pm$, $\sigma_{(1)}\chi_\pm = i\chi_\mp$. Из этих функций легко построить базисы двузначных представлений группы C_{3v}:

$$E_2': \quad \Psi_1 = \psi_+\chi_-, \ \Psi_2 = \psi_-\chi_+; \tag{5}$$

$$E_1': \quad \Psi_3 = 2^{-1/2}(\psi_+\chi_+ + \psi_-\chi_-), \ \Psi_4 = 2^{-1/2}(\psi_+\chi_+ - \psi_-\chi_-);$$

Очевидно, что матричные элементы между состояниями с различными χ могут возникнуть только от спин-орбитальной связи. Это же относится и к разности диагональных матричных элементов.

Функции (5) относятся к определенному $\mathbf{k}_i$. Точке $-\mathbf{k}_i$ соответствуют функции Ψ_1', Ψ_2', Ψ_3', Ψ_4', получаемые из функции (5) путем инверсии. Очевидно, что матричные элементы переходов между этими наборами могут возникнуть только в результате изменения периода и в первом приближении будут пропорциональны относительному сдвигу подрешеток.

3. Деформация, посредством которой из кубической решетки получается решетка типа висмута, может быть описана двумя величинами: вектором $\mathbf{u}$, направленным по пространственной диагонали куба и определяющим смещение подрешеток относительно центра инверсии, и компонентой тензора деформации $u_{x'y'} = u_{y'z'} = u_{z'x'}$ (координаты x', y', z' определены по отношению к кубическим осям). Отметим, что вектор $\mathbf{u}$ имеет разный знак для двух подрешеток, а потому, например, при преобразовании инверсии, когда происходит переход от одной подрешетки к другой, этот вектор должен считаться неизменным. Мы не учитываем здесь всестороннего сжатия, так как оно приводит лишь к смещению химического потенциала.

Для определения зависимости спектра от деформации мы должны составить инварианты из базисных функций, вектора $\boldsymbol{\varkappa} = \mathbf{k} - \mathbf{k}_i$, вектора $\mathbf{u}$ и тензора деформации u_{ik}. Мы будем считать, что компоненты $\boldsymbol{\varkappa}$, величины, определяющие деформации, а также величина Δ, характеризующая спин-орбитальную связь, — одного порядка величины (конечно с соответствующими коэффициентами) и ограничимся членами первого порядка. Отметим, что

из компонент u_{ik} составляется инвариант группы C_{3v}: $^1/_3 (u_{x'z'} + u_{y'z'} + u_{z'y'}) = w$. Примем во внимание, что функции Ψ'_i получаются из Ψ_i при инверсии.

С учетом этих замечаний напишем матрицу $D_{ss'}$. Легче всего сделать, используя систему отсчета с осью z по направлению $\mathbf{k}_i$. При этом получаем

$$D_{ss'} = \tag{6}$$

$$= \begin{Vmatrix}
f + a\varkappa_z + \Delta & 0 & \frac{b\varkappa_+ + \beta_+}{\sqrt{2}} & -\frac{b\varkappa_+ + \beta_+}{\sqrt{2}} & \gamma & 0 & \frac{\delta_+}{\sqrt{2}} & -\frac{\delta_+}{\sqrt{2}} \\
0 & f + a\varkappa_z + \Delta & \frac{b\varkappa_- + \beta_-}{\sqrt{2}} & \frac{b\varkappa_- + \beta_-}{\sqrt{2}} & 0 & \gamma & \frac{\delta_-}{\sqrt{2}} & \frac{\delta_-}{\sqrt{2}} \\
\frac{b\varkappa_- + \beta_-}{\sqrt{2}} & \frac{b\varkappa_+ + \beta_+}{\sqrt{2}} & f + a\varkappa_z - \Delta & 0 & \frac{\delta_-}{\sqrt{2}} & \frac{\delta_+}{\sqrt{2}} & \gamma & 0 \\
-\frac{b\varkappa_- + \beta_-}{\sqrt{2}} & \frac{b\varkappa_+ + \beta_+}{\sqrt{2}} & 0 & f + a\varkappa_z - \Delta & -\frac{\delta_-}{\sqrt{2}} & \frac{\delta_+}{\sqrt{2}} & 0 & \gamma \\
\gamma & 0 & \frac{\delta_+}{\sqrt{2}} & -\frac{\delta_+}{\sqrt{2}} & f - a\varkappa_z + \Delta & 0 & \frac{-b\varkappa_+ + \beta_+}{\sqrt{2}} & \frac{b\varkappa_+ - \beta_+}{\sqrt{2}} \\
0 & \gamma & \frac{\delta_-}{\sqrt{2}} & \frac{\delta_-}{\sqrt{2}} & 0 & f - a\varkappa_z + \Delta & \frac{-b\varkappa_- + \beta_-}{\sqrt{2}} & \frac{-b\varkappa_- + \beta_-}{\sqrt{2}} \\
\frac{\delta_-}{\sqrt{2}} & \frac{\delta_+}{\sqrt{2}} & \gamma & 0 & \frac{-b\varkappa_- + \beta_-}{\sqrt{2}} & \frac{-b\varkappa_+ + \beta_+}{\sqrt{2}} & f - a\varkappa_z - \Delta & 0 \\
-\frac{\delta_-}{\sqrt{2}} & \frac{\delta_+}{\sqrt{2}} & 0 & \gamma & \frac{b\varkappa_- - \beta_-}{\sqrt{2}} & \frac{-b\varkappa_+ + \beta_+}{\sqrt{2}} & 0 & f - a\varkappa_z - \Delta
\end{Vmatrix}$$

где

$$f = vu_z + \zeta w, \quad \gamma = cu_z, \quad \beta_\pm = eu_\pm = e(u_y \pm iu_x), \quad \delta_\pm = du_\pm. \tag{7}$$

Константы в матрице $D_{ss'}$ не зависят от того, окрестность какой точки $\mathbf{k}_i$ мы рассматриваем. Это следует из того, что полная величина $Q(\omega, \mathbf{r}, \mathbf{r}')$ должна быть инвариантом кубической группы (см. [7]).

В том случае, если мы рассматриваем окрестность точки $\mathbf{k}_0$, ось z направлена вдоль ромбоэдрической оси. Поэтому

$$f^{(0)} = vu + \zeta w, \quad \delta_\pm^{(0)} = \beta_\pm^{(0)} = 0, \quad \gamma^{(0)} = cu. \tag{8}$$

Если же $\mathbf{k} = \mathbf{k}_1$, то ось z направлена по другой оси C_3 в первоначальном кубическом кристалле. Преобразование, которое переводит вектор $\mathbf{k}_0$ в $\mathbf{k}_1$, есть поворот на π вокруг оси x' в кубической системе отсчета. Этот поворот надо совершать относительно оси, проходящей через узел кубической решетки, так как при этом не произойдет смены подрешеток. Проекция $\mathbf{u}$ на новую ось z равна $-u/3$. При повороте получаем

$$^1/_3 (u_{x'y'} + u_{x'z'} + u_{y'z'}) \to {}^1/_3 (u_{y'z'} - u_{x'y'} - u_{x'z'}) = -w/3.$$

Что же касается величин $u_\pm$, то их значение зависит от выбора осей. Удобно выбрать ось x в диагональной плоскости куба так, чтобы она была перпендикулярна кубической оси x'. После поворота новая ось x по-прежнему остается перпендикулярной оси x' и u_y окажется равным $-2\sqrt{2}\,u/3$. Таким образом, имеем

$$f^{(1)} = -vu/3 - \zeta w/3, \quad \beta_+^{(1)} = \beta_-^{(1)} = -2\sqrt{2}eu/3, \tag{9}$$

$$\gamma^{(1)} = -cu/3, \quad \delta_+^{(1)} = \delta_-^{(1)} = -2\sqrt{2}du/3.$$

Ниже будем для простоты применять обозначения

$$f^{(0)} = f, \quad \gamma^{(0)} = \gamma, \quad \beta^{(1)} = \beta, \quad \delta^{(1)} = \delta.$$

Для того чтобы определить спектр, надо найти собственные значения матрицы $D_{ss'}$. Мы будем предполагать, что $\beta_+ = \beta_- \neq 0$, $\delta_+ = \delta_- \neq 0$. Это дает нам возможность получить общую формулу для окрестностей $\mathbf{k}_0$ и $\mathbf{k}_1$.

Раскрывая $\det(-\omega\delta_{ss'} + D_{ss'})$, получаем уравнение для корней:

$$\Omega^4 + \Delta^4 + a^4\varkappa_z^4 + b^4(\varkappa_x^2 + \varkappa_y^2)^2 + \gamma^4 + \delta^4 - 2\Omega^2a^2\varkappa_z^2 - 2\Omega^2\Delta^2 - 2a^2\varkappa_z^2\Delta^2 - 2\Omega^2\gamma^2 + 2\gamma^2a^2\varkappa_z^2 - 2\gamma^2\Delta^2 + 2\Delta^2\beta^2 + 2\Delta^2b^2(\varkappa_x^2 + \varkappa_y^2) - 2\Omega^2\beta^2 - 2\Omega^2b^2(\varkappa_x^2 + \varkappa_y^2) - 2a^2\varkappa_z^2\beta^2 - 2a^2b^2(\varkappa_x^2 + \varkappa_y^2)\varkappa_z^2 + 2\delta^2a^2\varkappa_z^2 + 2\delta^2\Delta^2 - 2\delta^2\Omega^2 - 2\delta^2\beta^2 - 2\gamma^2\beta^2 + 2\gamma^2b^2(\varkappa_x^2 + \varkappa_y^2) - 2\gamma^2\delta^2 + 8(\gamma\delta - \Omega\beta)ab\varkappa_x\varkappa_y + 2(\delta^2 - \beta^2)b^2(\varkappa_y^2 - \varkappa_x^2) + 8\gamma\delta\beta\Omega = 0, \quad (10)$$

где $\Omega = f - \omega$; каждый корень является двукратным. Из-за членов, линейных по Ω, уравнение (10) в общем случае не может быть решено относительно Ω. Эти члены отсутствуют, если рассматривается окрестность $\mathbf{k}_0$, так как $\beta^{(0)} = \delta^{(0)} = 0$. При этом получаем

$$-\Omega = \omega - f = \pm\Big(\Delta^2 + a^2\varkappa_z^2 + \gamma^2 + b^2\varkappa_\perp^2 \pm 2\sqrt{\gamma^2\Delta^2 + a^2b^2\varkappa_z^2\varkappa_\perp^2 + \Delta^2a^2\varkappa_z^2}\Big)^{1/2}, \quad (11)$$

где $\varkappa_\perp^2 = \varkappa_x^2 + \varkappa_y^2$.

Рассмотрим, прежде всего, полученную формулу в пределе бесконечно малой деформации. Считая f, γ, β, δ стремящимися к нулю, получаем

$$\omega = \pm\,|\,|a\varkappa_z| \pm \sqrt{\Delta^2 + b^2\varkappa_\perp^2}\,|. \quad (12)$$

Нетрудно видеть, что изоэнергетическая поверхность $\omega = 0$ вырождается в двуполостный гиперболоид вращения. Этот гиперболоид является поверхностью, вдоль которой соприкасаются две функции $\omega(\mathbf{k})$. Для одной из них эта поверхность соответствует минимуму, а для другой максимуму. Это имеет место в окрестности любой из точек $\mathbf{k}_i$. Отсюда следует, что в пределе бесконечно малой деформации мы получаем вещество, промежуточное между металлом и диэлектриком, у которого полоса проводимости совершенно пуста, а валентная полоса заполнена; причем эти полосы касаются друг друга вдоль некоторой поверхности. Такое вырождение вдоль целой поверхности возможно только в предельном случае. Любая конечная деформация снимает это вырождение.

Теперь перейдем к конечной деформации. Согласно формуле (11), даже в этом случае можно найти значение ω, которое будет соответствовать минимуму одной из функций $\omega(\mathbf{k})$ в окрестности $\mathbf{k}_0$ и максимуму другой. Правда, при этом касание будет иметь место только в точках, а не вдоль поверхности. То же самое справедливо и для окрестности $\mathbf{k}_1$. Однако значения ω, при которых это имеет место в окрестности $\mathbf{k}_0$ и в окрестности $\mathbf{k}_1$, являются разными, что приведет к перекрытию зон. Это означает, что при любой конечной деформации в окрестности $\mathbf{k}_0$ и $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$ появятся поверхности Ферми, соответствующие электронам или дыркам.

В принципе формула (10) дает ответ на поставленную задачу. Остается только правильно выбрать химический потенциал. Для этого надо воспользоваться условием равенства числа электронов и числа дырок. Остальные коэффициенты, а именно a, b, Δ, γ, β, δ, f, находятся из экспериментальных данных (скажем, из эффекта де Гааз — ван Альфена и циклотронного резонанса). Отметим, что эти константы определяют весь электронный спектр, т. е. не только поверхность Ферми, но и любые производные $\partial^n\varepsilon/\partial p_i \ldots \partial p_k$ на границе. Полученные результаты дают возможность также изу-

чать влияние деформации на электронный спектр. При этом, конечно, надо знать не только тензор деформации, но и величину смещения подрешеток.

4. Для облегчения сопоставления результатов с экспериментом мы произведем анализ различных случаев. Начнем с окрестности $\mathbf{k} = \mathbf{k}_0$. Согласно экспериментальным данным, в окрестности $\mathbf{k} = \mathbf{k}_0$ находятся дырки, а в окрестности $\mathbf{k} = \mathbf{k}_1$ — электроны (знак носителей выбран условно). Ввиду этого в формуле (11) для окрестности $\mathbf{k}_0$ надо взять знак минус перед

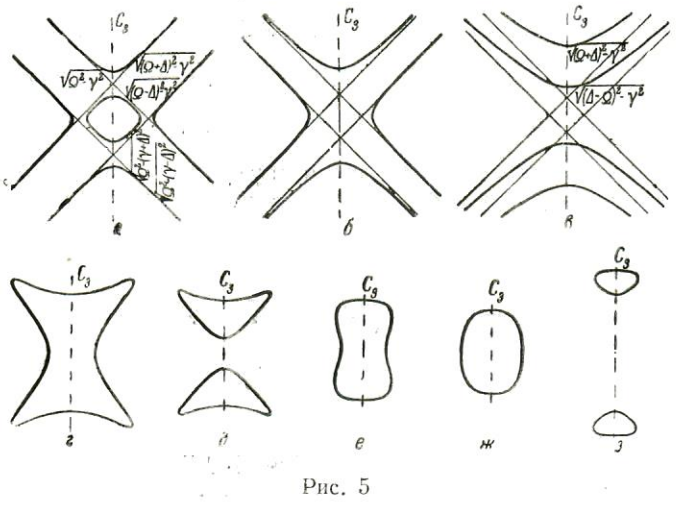


Рис. 5

корнем. Для упрощения формул мы введем обозначения $p = a\varkappa_z$ и $q = b\varkappa_\perp$. Кроме формулы (11), нам будет удобно решить уравнение (10) относительно p и относительно q. При этом получаем

$$p^2 = \Omega^2 + \Delta^2 - \gamma^2 + q^2 \pm 2\sqrt{\Omega^2\Delta^2 + (\Omega^2 - \gamma^2)\, q^2}, \qquad (13а)$$

$$q^2 = \Omega^2 + p^2 - \gamma^2 - \Delta^2 \pm 2\sqrt{\gamma^2\Delta^2 + (\Omega^2 - \gamma^2)\, p^2}. \qquad (13б)$$

Поскольку здесь входят только квадраты коэффициентов γ и Δ, то мы будем просто считать их положительными. Положительность Ω следует из (11).

Из формулы (13) сразу следуют два принципиально различных случая. Если $\Omega > \gamma$, то p и q могут принимать бесконечные значения, если же $\Omega < \gamma$, то p и q должны оставаться конечными. В действительности, конечно, все наше рассмотрение пригодно до тех пор, пока $\varkappa_\perp$ и $\varkappa_z$ заметно меньше периода обратной решетки. Более детальный анализ показывает, что в зависимости от соотношений коэффициентов Ω, γ, Δ возникает несколько характерных типов изоэнергетических поверхностей. Все эти типы изображены на рис. 5.

Открытые поверхности возникают в случае $\Omega > \gamma$. Изображенные типы поверхностей соответствуют

$$а — \Omega > \Delta + \gamma;\ б — \gamma < \Omega < \gamma + \Delta,\ \Omega > \Delta - \gamma;\ в — \gamma < \Omega < \Delta - \gamma.$$

Отметим, что в случае $а$ имеется две ферми-поверхности — замкнутая и открытая (мы их изобразили вместе). Внутри каждой из них располагается область больших энергий.

Замкнутые ферми-поверхности возникают в случаях

$$е — \gamma > \Omega > \sqrt{\Delta^2/4 + \gamma^2} - \Delta/2,\ \Omega > \Delta - \gamma;$$
$$д — \gamma > \Omega > \sqrt{\Delta^2/4 + \gamma^2} - \Delta/2,\ \Omega < \Delta - \gamma;$$

$$е — \sqrt{\Delta^2/4 + \gamma^2} - \Delta/2 > \Omega > \sqrt{\gamma(\gamma - \Delta)},\ \gamma > \Delta;$$

$$ж — \sqrt{\gamma(\gamma - \Delta)} > \Omega > \gamma - \Delta,\ \gamma > \Delta;$$

$$з — \sqrt{\Delta^2/4 + \gamma^2} - \Delta/2 > \Omega > 0,\ \Omega < \Delta - \gamma.$$

Заметим, что замкнутые поверхности получаются только в том случае, когда $\Delta \neq 0$, т. е. благодаря спин-орбитальной связи. Детальное сравнение с экспериментальными данными будет произведено в следующей работе.

Теперь рассмотрим положение с электронами. В окрестности точек $\mathbf{k}_1$ зависимость $\omega(\mathbf{k})$ имеет более низкую симметрию C_{2h} и зависит от двух дополнительных параметров, β и δ. Формула (10) позволяет в этом случае выразить $\varkappa_x$, как функцию $\varkappa_y$, $\varkappa_z$ и Ω. Введя координаты $q_x = b\varkappa_x$, $q_y = b\varkappa_y$, находим

$$q_x^2 = p^2 - q_y^2 + \Omega^2 - \gamma^2 + \delta^2 - \beta^2 - \Delta^2 \pm$$
$$\pm 2\sqrt{(\beta q_y + \Omega p)^2 - (\delta q_y + \gamma p)^2 + (\Omega\delta - \gamma\beta)^2 + \Delta^2(\gamma^2 - \delta^2)}. \quad (14)$$

Произвести полный анализ всех возможностей, вытекающих из этой формулы, затруднительно. Мы определим только, при каких условиях поверхность $\omega = \text{const}$ будет замкнутой. Для этого рассмотрим соотношение (14) при $\varkappa \to \infty$. Отбрасывая постоянные члены, получаем

$$q_x^2 \approx p^2 - q_y^2 \pm 2\sqrt{(\beta q_y + \Omega p)^2 - (\delta q_y + \gamma p)^2}. \quad (15)$$

Основные члены здесь квадратичные. Очевидно, что в направлении, соответствующем $q_y^2 > p^2$, поверхность будет конечной. Для того чтобы это оставалось правильным при всех направлениях, надо, чтобы при $q_y^2 < p^2$ выражение под корнем было бы отрицательным, иначе говоря

$$(\beta s + \Omega)^2 - (\delta s + \gamma)^2 < 0,\ s^2 < 1, \quad (16)$$

где $s = q_y/p$. Записав это неравенство в виде

$$[(\beta + \delta)s + (\gamma + \Omega)]\,[(\beta - \delta)s - (\gamma - \Omega)] < 0,$$

мы немедленно получаем условие на коэффициенты:

$$|\gamma + \Omega| > |\beta + \delta|,\ |\gamma - \Omega| > |\beta - \delta|,\ \gamma^2 - \Omega^2 > 0. \quad (17)$$

5. Для полного объяснения особенностей висмута остается решить два вопроса. Во-первых, надо доказать, что полученный спектр устойчив по отношению к кулоновскому взаимодействию квазичастиц. Во-вторых, надо продемонстрировать, что фазовый переход металлической фазы в «диэлектрическую» действительно может иметь место.

Первый из этих вопросов решается легко. Поскольку мы использовали линейное приближение по $\varkappa$ в энергии, то скорость на ферми-границе будет порядка обычных электронных скоростей даже при малых $\varkappa$. Это могло бы быть неверным, если бы $a\varkappa_z$ или $b\varkappa_\perp$ была гораздо меньше γ, но условие равенства чисел электронов и дырок и тот факт, что $f \sim \gamma$, исключает такую возможность. Согласно предыдущему, при этом радиус экранировки имеет порядок расстояния между квазичастицами.

Второй вопрос значительно сложнее, и мы не можем полностью его решить. Основная трудность заключается в невозможности определения энергии металлической фазы. Однако энергию диэлектрической фазы можно найти. Изменение энергии при переходе от кубической решетки к решетке висмута будет состоять из изменения энергии ионов, которое, очевидно, пропорционально квадрату деформации, и изменения энергии электронов, включая взаимодействие электронов с ионами. Вот этой-то энергией мы и займемся.

Ввиду сложности спектра в общем случае мы упростим себе задачу, положив $\beta = \delta = 0$. При этом в окрестности $\mathbf{k}_1$ спектр будет тем же, что и в окрестности $\mathbf{k}_0$, но только с другими коэффициентами.

Мы рассмотрим два случая: $f, \gamma \ll \Delta$ и $f, \gamma \gg \Delta$. Начнем со случая $f, \gamma \ll \Delta$. Для нахождения электронной энергии, очевидно, надо вычислить $\int_0^{N_0} \mu(N) dN$, где N_0 соответствует реальному числу электронов. Интеграл этот можно записать в виде

$$\mu N_0 - \int_{\mu(0)}^{\mu} N(\mu_1)\, d\mu_1, \tag{18}$$

где $\mu = \mu(N_0)$ — фермиевская энергия.

Прежде всего определим величину μ в нашей модели. Для этого найдем значение ω, при котором объемы «электронной» и «дырочной» поверхностей будут равны. Будем считать, что это осуществляется при замкнутых ферми-поверхностях, т. е. при наших предположениях, в случае ∂ или $з$. Согласно формуле (13), объем дырочной ферми-поверхности в этом случае пропорционален

$$\int_0^{q_{max}} qdq \{[\Omega^2 + \Delta^2 - \gamma^2 + q^2 + 2\sqrt{\Omega^2\Delta^2 + (\Omega^2 - \gamma^2)\, q^2}]^{1/2} -$$
$$- [\Omega^2 + \Delta^2 - \gamma^2 + q^2 - 2\sqrt{\Omega^2\Delta^2 + (\Omega^2 - \gamma^2)\, q^2}]^{1/2}\} \approx \tag{19}$$
$$\approx 2 \int_0^{q_{max}} y^2 dy \left(\frac{\Delta^2\gamma^2}{\gamma^2 - \Omega^2} - y^2\right)^{-1/2} = \Delta^2 \left(-|\Omega| + \frac{\gamma^2}{\sqrt{\gamma^2 - \Omega^2}} \arcsin \frac{|\Omega|}{|\gamma|}\right),$$

где $q_{max} = |\Omega|\, \Delta / \sqrt{\gamma^2 - \Omega^2}$ (здесь мы пренебрегли членами порядка $(\Omega/\Delta)^2$, $(\gamma/\Delta)^2$).

Для дырок в качестве Ω надо подставить $f - \mu$; это величина положительная. Для электронов в выражении для ω надо брать корень со знаком плюс. Следовательно, $\Omega = f^{(1)} - \mu < 0$, т. е. $|\Omega| = \mu - f^{(1)}$. Согласно формуле (9), $\gamma^{(1)} = -\gamma/3$, $f^{(1)} = -f/3$. Кроме того, надо учесть, что всего имеется три неэквивалентных точки $\mathbf{k}_1$, образующие звезду, т. е. электронный объем должен быть утроен. С учетом этих замечаний легко увидеть, что условие равенства объемов поверхностей электронов и дырок дает

$$\mu = 0.$$

Теперь перейдем к вычислению электронной энергии. Поскольку нас интересует только изменение энергии при деформации, то из выражения (18) надо вычесть соответствующие выражения для недеформированного кристалла. При этом получаем

$$-\int_{\mu(0)}^{0} [N(f, \gamma, \mu) - N(0, 0, \mu)]\, d\mu.$$

Как уже отмечалось, в пределе, когда деформация уменьшается до нуля, минимум верхней энергетической полосы совпадает с максимумом нижней. При этом верхняя полоса будет пуста, а нижняя — заполнена. Отсюда вытекает, что

$$N(f, \gamma, \mu) - N(0, 0, \mu) = N_э(f, \gamma, \mu) - N_д(f, \gamma, \mu) + N_д(0, 0, \mu),$$

где $N_э$ и $N_д$ — числа электронов и дырок, а $N_д(0, 0, 0) = 0$ (здесь подразумевается $\mu < 0$).

В окрестности $\mathbf{k} = \mathbf{k}_0$ при $\mu < 0$ мы всегда имеем дырки, а в окрестности $\mathbf{k} = \mathbf{k}_1$ при $\mu > -f/3$ мы имеем электроны, а при $\mu < -f/3$ — дырки. Таким образом, искомый интеграл будет равен

$$\int\limits_{\mu(0)}^{0} N_{д0}(f, \gamma, \mu)\, d\mu + 3\int\limits_{\mu(0)}^{-f/3} N_{д1}(f, \gamma, \mu)\, d\mu - 3\int\limits_{-f/3}^{0} N_{э1}(f, \gamma, \mu)\, d\mu - \\ - 4\int\limits_{\mu(0)}^{0} N_{д}(0, 0, \mu)\, d\mu,$$

где $N_{д0}$ соответствует дыркам в окрестности $\mathbf{k}_0$, а $N_{д1}$ — в окрестности $\mathbf{k}_1$. Коэффициент 3 появляется от суммирования по всем точкам звезды.

Поскольку, как мы видели, объемы в действительности зависят только от $|\Omega|$, то мы произведем замену переменных и получим

$$E_э = \int\limits_{f}^{-\mu(0)+f} d|\Omega|\, N_д(\gamma, |\Omega|) + 3\int\limits_{f/3}^{-\mu(0)-f/3} d|\Omega|\, N_д\left(-\frac{\gamma}{3}, |\Omega|\right) - \\ - 4\int\limits_{0}^{-\mu(0)} N_д(0, |\Omega|)\, d|\Omega|. \qquad (20)$$

Числа дырок пропорциональны соответствующим объемам поверхности с заданным Ω. Очевидно, основная часть интегралов будет соответствовать значениям $|\Omega| > \gamma$, где поверхности являются незамкнутыми. Объем, ограниченный какой-нибудь из этих поверхностей, согласно (13), пропорционален

$$N_д \sim \int\limits_{0}^{\infty} [p_{max}(q) - p_{min}(q)]\, qdq = \\ = \int\limits_{0}^{\infty} \{[\Omega^2 + \Delta^2 - \gamma^2 + q^2 + 2\sqrt{\Omega^2\Delta^2 + (\Omega^2 - \gamma^2)\, q^2}\,]^{1/2} - \\ - [\Omega^2 + \Delta^2 - \gamma^2 + q^2 - 2\sqrt{\Omega^2\Delta^2 + (\Omega^2 - \gamma^2)\, q^2}\,]^{1/2}\}\, qdq.$$

В таком виде эта формула, строго говоря, годится только для $\Omega < \Delta + \gamma$, но, как мы сейчас увидим, в интеграле по q в основном существенны большие значения q, а потому такая формула оказывается справедливой для произвольных Ω.

Считая q большим и производя замену переменных, получаем

$$N_д \sim \int\limits_{\frac{\Omega\Delta}{\sqrt{\Omega^2-\gamma^2}}}^{L} y^2 dy \sqrt{\Omega^2 - \gamma^2}\left(y^2 - \frac{\gamma^2\Delta^2}{\Omega^2 - \gamma^2} + \Omega^2 - \gamma^2\right)^{-1/2} \approx L^2\sqrt{\Omega^2 - \gamma^2}. \qquad (21)$$

В качестве верхнего предела здесь взято значение квазиимпульса q в области, где полученные формулы для энергетического спектра перестают соответствовать действительности. Очевидно, $L \sim bK$, где K — величина порядка периода обратной решетки. Отметим тут же, что в случае $f, \gamma \gg \Delta$ равенство чисел электронов и дырок может быть достигнуто только с незамкнутыми поверхностями. При этом соответствующие объемы выражаются формулой (21). Если брать одно и то же L для электронов и дырок (это очевидно следует из того факта, что $L \gg \gamma, f$), то условие равенства чисел электронов и дырок опять дает $\mu = 0$. Это следует из того, что $|\gamma^{(1)}| = \gamma/3$, $|f^{(1)}| = f/3$, и число электронных поверхностей равно трем.

Подставляя формулу (21) в (20), получаем

$$E_э \sim L^2 \left\{ \int\limits_{\max(f,\gamma)}^{-\mu(0)} d|\Omega| \sqrt{\Omega^2 - \gamma^2} + 3 \int\limits_{\max\left(\frac{f}{3},\frac{\gamma}{3}\right)}^{-\mu(0)} d|\Omega| \sqrt{\Omega^2 - \frac{\gamma^2}{9}} - \right.$$

$$\left. - 4 \int\limits_0^{-\mu(0)} \Omega d\Omega \approx -\frac{2}{3} L^2 \gamma^2 \ln \frac{|\mu(0)|}{\max(f,\gamma)} \right\}. \qquad (22)$$

Замечательным свойством этой энергии является то, что при малых деформациях, она может стать больше любого члена квадратичного по γ.

Следует отметить, что произведенный вывод не является вполне строгим. Ввиду того что в интеграле по $\varkappa$ существенны больше $\varkappa$, необходимо было бы учесть квадратичные члены. Однако, ввиду того что в «диэлектрической» фазе при $\gamma \to 0$ имеем [3] $\Omega \ll |\mu(0)|$ и $N_д \sim \Omega$, мы, очевидно, обязательно получим при $\gamma \ll \Omega \ll |\mu(0)|$, что $N_д = c_1\Omega + c_2\gamma^2/\Omega$ (c_1, c_2 — константы), откуда следует (22). Знак коэффициента c_2, вообще говоря, может быть любым. По-видимому, его следует считать отрицательным, так как максимальные значения компонент $\varkappa$ все же в несколько раз меньше первоначального периода обратной решетки, и поэтому произведенный вывод дает правильное указание относительно знака энергии.

Добавляя энергию ионов, находим энергию деформации кубической решетки для диэлектрической фазы (очевидно $f \sim \gamma$):

$$E_{диэл}(\gamma) = E_{диэл}(0) + \sigma\gamma^2 - \chi\gamma^2 \ln \frac{P}{\gamma} \qquad (23)$$

($P = \text{const} \approx 1$ eV; $\sigma, \chi > 0$). Минимизация этой энергии по γ дает

$$\gamma_0 = Pe^{-(2\sigma+\chi)/2\chi}. \qquad (24)$$

Представляется вероятным, что константа σ, определяющая энергию взаимодействия ионов, $\sigma\gamma^2$, несколько больше константы χ, соответствующей взаимодействию электронов с измененным потенциалом. При этом из (24) следует $\gamma \ll P$, т. е. малость деформации. Подставляя (24) в (23), получаем

$$E_{диэл} = E_{диэл}(0) - \chi\gamma_0^2/2. \qquad (25)$$

Для металлической фазы расчет энергии произвести не удается. Если эта энергия растет при увеличении деформации, причем

$$E_{диэл}(0) - E_{мет}(0) < \chi\gamma_0^2/2, \qquad (26)$$

то переход из кубической решетки в решетку висмута является энергетически выгодным. Малость щели между двумя фазами означает, что висмут сравнительно легко может быть переведен в металлическое состояние, что действительно наблюдается на опыте [10].

В заключение авторы выражают благодарность И. М. Лифшицу, И. М. Халатникову, Л. П. Горькову и И. Е. Дзялошинскому за обсуждение работы.

Институт физических проблем
Академии наук СССР

Поступила в редакцию
23 апреля 1962 г.

[3] Это связано с тем, что при $\gamma \to 0$ полоса проводимости касается валентной полосы вдоль поверхности.

Литература

[1] D. Schoenberg. Proc. Roy. Soc., **A170**, 341, 1939; Phil. Trans. Roy. Soc. **A245**, 1, 1952. Н. Б. Брандт, М. В. Разуменко. ЖЭТФ, **39**, 276, 1960. Н. Б. Брандт, В. В. Щекочихина. ЖЭТФ, **41**, 1412, 1961.
[2] Н. Б. Брандт, А. Е. Дубровская, Г. А. Кытин. ЖЭТФ, **37**, 572. 1959. Н. Б. Брандт. ПТЭ, 2, 1960. Н. Б. Брандт. ЖЭТФ, **38**, 1355, 1960.
[3] J. M. Luttinger. Phys. Rev., **119**, 1153, 1960.
[4] B. S. Boyle, K. F. Rodgers. Phys. Rev. Lett., **2**, 338, 1959.
[5] B. Lax. Bull. Am. Phys. Soc., **5**, 167, 1960.
[6] H. Johns. Proc. Roy. Soc., **A144**, 225, 1934; **A147**, 396, 1934.
[7] А. А. Абрикосов. ЖЭТФ, этот выпуск, стр. 1083.
[8] Л. Д. Ландау, Е. М. Лифшиц. Квантовая механика, Гостехиздат, 1948.
[9] R. Elliot. Phys. Rev., **96**, 280, 1954.
[10] Н. Б. Брандт, Н. И. Гинзбург. ЖЭТФ, **39**, 1554, 1960.

THEORY OF THE ELECTRON ENERGY SPECTRUM OF METALS WITH A BISMUTH TYPE LATTICE

A. A. Abrikosov, L. A. Falkovsky

The electron energy spectrum of metals of the fifth group of the periodic system is determined on basis of some qualitative assumptions and symmetry theory.