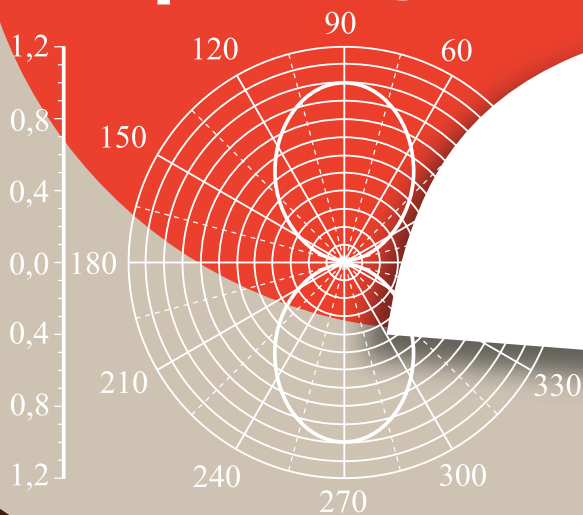


Н. П. Калашников  
М. А. Смондырев

# ОСНОВЫ ФИЗИКИ



УЧЕБНИК ДЛѦ ВЫСШЕЙ ШКОЛЫ

Н. П. Калашников, М. А. Смондырев

# ОСНОВЫ ФИЗИКИ

## Том 2

2-Е ИЗДАНИЕ, ЭЛЕКТРОННОЕ



Москва  
Лаборатория знаний  
2021

УДК 53(075.8)  
ББК 22.3я73  
К17

*Серия основана в 2009 г.*

**Калашников Н. П.**

К17 Основы физики : в 3 т. Т. 2 / Н. П. Калашников, М. А. Смондырев. — 2-е изд., электрон. — М. : Лаборатория знаний, 2021. — 609 с. — (Учебник для высшей школы). — Систем. требования: Adobe Reader XI ; экран 10". — Загл. с титул. экрана. — Текст : электронный.

ISBN 978-5-00101-075-3 (Т. 2)

ISBN 978-5-00101-072-2

Учебник соответствует программе дисциплины «Физика» для естественнонаучных и технических университетов. Два его тома входят в состав учебного комплекта, включающего также учебное пособие «Основы физики. Упражнения и задачи» тех же авторов.

Во многих отношениях данный учебник не имеет аналогов. Ряд оригинальных методических приемов и способов изложения материала, включение новых, зачастую неожиданных тем и ярких примеров, отсутствующих в традиционных курсах физики, позволяют учащимся приобрести навыки уверенного самостоятельного мышления, глубже понять физические основы самых различных природных явлений, делать практические, качественные оценки, оперируя размерностями и порядками величин.

Для студентов естественнонаучных и инженерно-технических специальностей.

**УДК 53(075.8)  
ББК 22.3я73**

**Деривативное издание на основе печатного аналога:** Основы физики : в 3 т. Т. 2 / Н. П. Калашников, М. А. Смондырев. — М. : Лаборатория знаний, 2017. — 606 с. : ил. — (Учебник для высшей школы). — ISBN 978-5-00101-005-0 (Т. 2); ISBN 978-5-00101-003-6.

**В соответствии со ст. 1299 и 1301 ГК РФ при устранении ограничений, установленных техническими средствами защиты авторских прав, правообладатель вправе требовать от нарушителя возмещения убытков или выплаты компенсации**

ISBN 978-5-00101-075-3 (Т. 2)  
ISBN 978-5-00101-072-2

© Лаборатория знаний, 2017

## Часть IV

# Колебания и волны

*И тогда я увидел Маятник.  
Движущийся шар на конце длинной нити,  
прикрепленной к своду центрального  
купола, описывал широкие колебательные  
движения ритмично и величаво.  
Любой поддался бы очарованию этого  
мирного дыхания.*

---

У. Эко. «Маятник Фуко»



## Глава 24

# Колебательное движение

Физические процессы, характеризующиеся той или иной степенью повторяемости, называются колебаниями. В зависимости от физической природы повторяющегося процесса различают колебания механические, электромагнитные, электромеханические и т. д.

В зависимости от характера воздействия на колеблющееся тело различают свободные (или собственные) и вынужденные колебания.

Если положение системы в любое время может быть описано единственным параметром, то система имеет одну степень свободы. Примеры таких систем: маятник, колеблющийся в заданной плоскости; масса, связанная с пружиной;  $LC$ -цепочка (рис. 24.1). Обычно для таких систем употребляют общее название *осциллятор* (от англ. *oscillate* — колебаться, вибрировать).

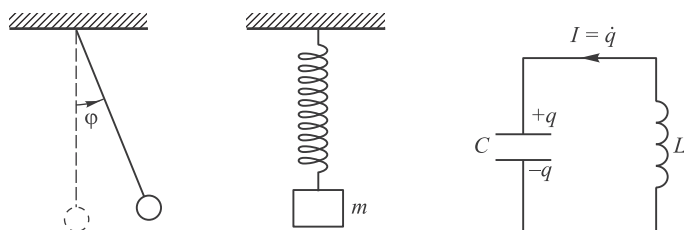


Рис. 24.1. Примеры колебательных систем с одной степенью свободы

### 24.1 Уравнение гармонических колебаний

В этом разделе мы покажем, что уравнения колебательного движения многих систем в сущности одинаковы, так что различные физические процессы могут быть описаны одними и теми же математическими формулами.

#### Пружинный маятник

Рассмотрим систему, состоящую из шарика массой  $m$ , подвешенного на пружине (рис. 24.2). В положении равновесия сила тяжести  $mg$  уравновешивается упругой

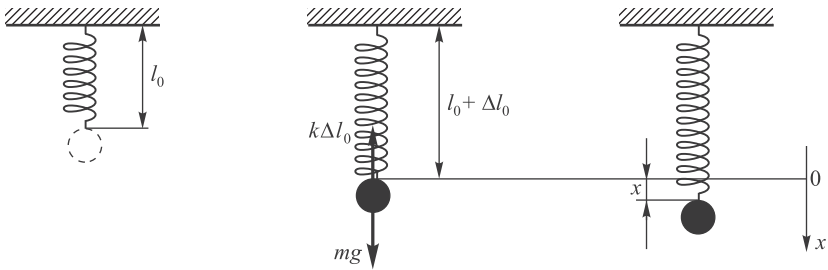


Рис. 24.2. Пружинный маятник (к выводу уравнения движения)

силой  $k\Delta l_0$ :

$$mg = k\Delta l_0, \quad \text{откуда} \quad \Delta l_0 = \frac{mg}{k}, \quad (24.1)$$

где  $\Delta l_0$  — статическое удлинение пружины. Направим ось  $x$  вниз и выберем начало отсчета так, чтобы координата  $x = 0$  соответствовала положению неподвижного шарика в положении равновесия.

Если теперь оттянуть шарик от положения равновесия на расстояние  $x$ , то полное удлинение пружины станет равным  $\Delta l_0 + x$ . С учетом закона Гука результирующая сила, действующая на шарик, будет тогда равна:

$$F = mg - k(\Delta l_0 + x). \quad (24.2)$$

Используем соотношение (24.1) и в результате получим, что

$$F = -kx. \quad (24.3)$$

Знак «минус» означает, что сила стремится уменьшить отклонение шарика от положения равновесия. Полученное выражение соответствует упругой силе слабо деформированной пружины.

Запишем теперь уравнение второго закона Ньютона:  $m\ddot{x} = -kx$ . Его можно также представить в виде

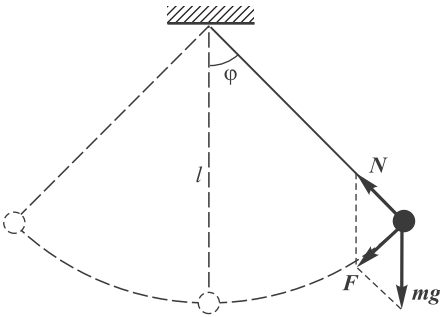
$$\ddot{x} + \omega_0^2 x = 0, \quad \omega_0 = \sqrt{\frac{k}{m}}. \quad (24.4)$$

## Математический маятник

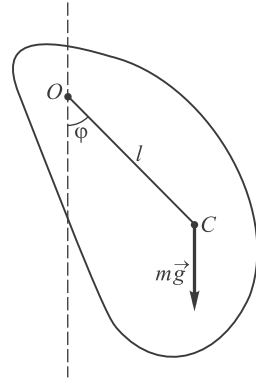
*Математический маятник* представляет собой идеализированную систему, состоящую из невесомой и нерастяжимой нити, на которой подвешена масса, сосредоточенная в одной точке.

Будем характеризовать отклонение маятника от положения равновесия углом  $\varphi$ , который образует нить с вертикалью (рис. 24.3). При отклонении маятника от положения равновесия на материальную точку массой  $m$  действуют сила тяжести  $mg$  и сила натяжения нити  $N$ . Их равнодействующая  $F$  направлена по касательной к окружности радиусом  $l$  и равна  $F = -mg \sin \varphi$ . Скорость материальной точки тоже направлена по касательной и равна  $v = l\dot{\varphi}$ , так что тангенциальное ускорение  $a_\tau = \dot{v} = l\ddot{\varphi}$ . Записываем теперь уравнение движения:

$$ml\ddot{\varphi} = -mg \sin \varphi \quad (24.5)$$



**Рис. 24.3.** Силы, возникающие при колебаниях математического маятника (к выводу уравнения движения)



**Рис. 24.4.** При выводе уравнения движения физического маятника необходимо учитывать его момент инерции  $J$

(знак «минус» соответствует тому, что сила  $F$  стремится уменьшить угол  $\varphi$ ). При небольших отклонениях маятника  $\sin \varphi \approx \varphi$ . Получаем

$$\ddot{\varphi} + \omega_0^2 \varphi = 0, \quad \omega_0 = \sqrt{\frac{g}{l}}. \quad (24.6)$$

### Физический маятник

Если колеблющееся тело, подвешенное на оси, невозможно представить как материальную точку, маятник называется *физическим* (рис. 24.4). При отклонении маятника от положения равновесия на угол  $\varphi$  возникает вращательный момент, стремящийся вернуть маятник в положение равновесия. Этот момент равен:

$$M = -mgl \sin \varphi, \quad (24.7)$$

где  $m$  — масса маятника, а  $l$  — расстояние  $OC$  между точкой подвеса  $O$  и центром масс  $C$  маятника.

Рассматривая  $\varphi$  как вектор, связанный с направлением поворота правилом правого винта, противоположность знаков  $M$  и  $\varphi$  можно объяснить тем, что векторы  $\vec{M}$  и  $\vec{\varphi}$  направлены в противоположные стороны. Обозначив момент инерции маятника относительно оси, проходящей через точку подвеса, через  $J$ , можно записать основное уравнение динамики вращательного движения:

$$J\ddot{\varphi} = -mgl \sin \varphi. \quad (24.8)$$

Ограничимся рассмотрением малых колебаний:  $\sin \varphi \approx \varphi$ . В этом случае уравнение колебаний принимает вид

$$\ddot{\varphi} + \omega_0^2 \varphi = 0, \quad \omega_0 = \sqrt{\frac{mgl}{J}}. \quad (24.9)$$

В случае, когда физический маятник можно представить как материальную точку, колеблющуюся на нити длиной  $l$ , момент инерции равен  $J = ml^2$ , и мы приходим к уравнению (24.6) движения математического маятника.



## Движение поршня в сосуде с идеальным газом

Рассмотрим поршень массой  $m$  и площадью поверхности  $S$ , закрывающий сосуд объемом  $V_0$  с идеальным газом, изолированным от окружающей среды (рис. 24.5). Пусть в состоянии равновесия давление в сосуде равно  $p_0$ . Это давление складывается из атмосферного давления  $p_a$  и давления  $mg/S$ , оказываемого поршнем:

$$p_0 = p_a + \frac{mg}{S}. \quad (24.10)$$

Переместим поршень на расстояние  $x$ . Объем сосуда увеличится и станет равным  $V = V_0 + Sx$ . Соответственно уменьшится давление. Поскольку процесс идет без теплообмена, применяем уравнение адиабаты (см. т. 1, уравнение (11.30)) и находим новое давление:

$$pV^\gamma = p_0V_0^\gamma, \quad \text{откуда} \quad p = \frac{p_0}{(1 + Sx/V_0)^\gamma}. \quad (24.11)$$

Здесь  $\gamma$  — показатель адиабаты, зависящий от числа степеней свободы молекул газа. При малых колебаниях, когда смещения поршня много меньше высоты сосуда ( $x \ll V_0/S$ ), можно разложить  $p$  в ряд Тейлора:

$$p \approx p_0 \left( 1 - x \frac{\gamma S}{V_0} \right). \quad (24.12)$$

На поршень действуют три силы: сила атмосферного давления  $-p_a S$ , сила давления газа в сосуде  $pS$  и сила тяжести  $-mg$ . Знаки сил соответствуют выбору положительного направления оси  $x$  вверх. Используя (24.10) и (24.12), находим для равнодействующей  $F$  этих сил:

$$\begin{aligned} F &= -p_a S + pS - mg = -S \left( p_0 - \frac{mg}{S} \right) + Sp_0 \left( 1 - x \frac{\gamma S}{V_0} \right) - mg = \\ &= -x \frac{\gamma p_0 S^2}{V_0}. \end{aligned} \quad (24.13)$$

Записываем теперь уравнение движения поршня  $m\ddot{x} = F$  в виде

$$\ddot{x} + \omega_0^2 x = 0, \quad \omega_0 = \sqrt{\frac{\gamma p_0 S^2}{mV_0}}. \quad (24.14)$$

## Электромагнитный контур

Рассмотрим колебательный контур, состоящий из конденсатора емкостью  $C$  и катушки индуктивностью  $L$  (рис. 24.6). Сопротивлением катушки и проводов пренебрегаем. Пусть в цепи идет ток  $I$ , заряжающий конденсатор:  $I = dq/dt$ . Так как внешняя ЭДС к контуру не приложена, ЭДС самоиндукции  $\mathcal{E} = -L\dot{I}$  равна напряжению  $q/C$  на конденсаторе. Имеем два уравнения:

$$I = \frac{dq}{dt}, \quad L \frac{dI}{dt} + \frac{q}{C} = 0. \quad (24.15)$$

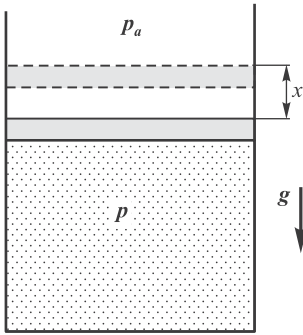


Рис. 24.5. Колебания поршня, закрывающего сосуд с идеальным газом

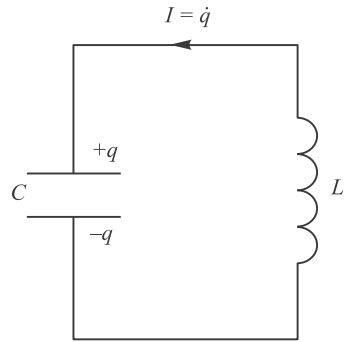


Рис. 24.6. Электромагнитный колебательный контур

Подставляя первое уравнение во второе, получаем уравнение для изменения заряда на конденсаторе:

$$\ddot{q} + \omega_0^2 q = 0, \quad \omega_0 = \frac{1}{\sqrt{LC}}. \quad (24.16)$$

Вместо использованной подстановки выражения тока через заряд можно продифференцировать второе из уравнений (24.15) и выразить производную от заряда через ток. В результате получим аналогичное уравнение для изменения тока в цепи с тем же выражением для  $\omega_0$ , что и в (24.16):

$$\ddot{I} + \omega_0^2 I = 0. \quad (24.17)$$

## 24.2 Гармонические колебания

Мы рассмотрели несколько совершенно различных систем и убедились, что уравнения движения приводятся к одной и той же форме:

$$\ddot{x} + \omega_0^2 x = 0. \quad (24.18)$$

Разница между физическими системами сводится к различию в формулах для частоты собственных колебаний  $\omega_0$ , что связано с разным физическим смыслом переменной  $x$ : это может быть координата, угол, заряд, ток и т. д. Уравнение (24.18) описывает так называемые **гармонические** колебания. Гармоническими колебаниями называют такие колебательные движения, при которых смещение тела от положения равновесия совершается по закону синуса или косинуса. Такие колебания иногда называют *малыми колебаниями*, потому что линейный закон для «возвращающей силы» (в широком смысле слова) характерен именно для малых отклонений системы от положения равновесия.

Уравнение гармонических колебаний (24.18) является линейным дифференциальным уравнением второго порядка (так как оно содержит вторую производную от переменной  $x$ ). Линейность уравнения означает, что:

- 1) если какая-то функция  $x(t)$  является решением этого уравнения, то функция  $Cx(t)$  также будет его решением ( $C$  — произвольная постоянная);
- 2) если функции  $x_1(t)$  и  $x_2(t)$  являются решениями этого уравнения, то их сумма  $x_1(t) + x_2(t)$  также будет решением того же уравнения.

Доказана также математическая теорема, что уравнение второго порядка имеет два независимых решения. Все остальные решения, согласно свойствам линейности, могут быть получены как их линейные комбинации.

Непосредственным дифференцированием легко проверить, что независимые функции  $\sin \omega_0 t$  и  $\cos \omega_0 t$  удовлетворяют уравнению (24.18). Значит, общее решение этого уравнения имеет вид

$$x(t) = C_1 \sin \omega_0 t + C_2 \cos \omega_0 t, \quad (24.19)$$

где  $C_1$  и  $C_2$  — произвольные постоянные. Это решение может быть представлено и в другом виде. Введем величину

$$A = \sqrt{C_1^2 + C_2^2} \quad (24.20)$$

и определим угол  $\alpha$  соотношениями

$$\cos \alpha = \frac{C_2}{A} = \frac{C_2}{\sqrt{C_1^2 + C_2^2}}, \quad \sin \alpha = -\frac{C_1}{A} = -\frac{C_1}{\sqrt{C_1^2 + C_2^2}}. \quad (24.21)$$

Тогда общее решение (24.19) можно записать как

$$x(t) = A (\cos \omega_0 t \cos \alpha - \sin \omega_0 t \sin \alpha). \quad (24.22)$$

Согласно формулам тригонометрии, выражение в скобках равно  $\cos(\omega_0 t + \alpha)$ . Окончательно приходим к общему решению уравнения гармонических колебаний в виде

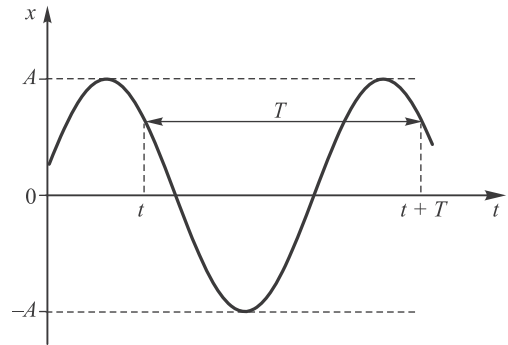
$$\boxed{x(t) = A \cos(\omega_0 t + \alpha)}. \quad (24.23)$$

Величина  $A$  называется **амплитудой колебания**, а  $\alpha$  — **начальной фазой**. Вся комбинация  $\omega_0 t + \alpha$  называется **фазой колебания**. Выражения (24.19) и (24.23) совершенно эквивалентны, так что мы для простоты можем пользоваться любым из них. Оба решения являются периодическими функциями времени. Действительно, синус и косинус периодичны с периодом  $2\pi$ . Поэтому различные состояния системы, совершающей гармонические колебания, повторяются через промежуток времени  $t^*$ , за который фаза колебания получает приращение, кратное  $2\pi$ :

$$\omega_0(t + t^*) + \alpha = \omega_0 t + \alpha + 2\pi n, \quad n = 0, \pm 1, \pm 2, \dots \quad (24.24)$$

Отсюда следует, что  $t^* = 2\pi n / \omega_0$ . Наименьшее из этих значений

$$\boxed{T = \frac{2\pi}{\omega_0}} \quad (24.25)$$



**Рис. 24.7.** Временная развертка гармонического колебания с периодом  $T$

называется **периодом колебаний** (рис. 24.7), а  $\omega_0$  — **круговой (циклической) частотой**. Равномерное движение по окружности также можно рассматривать как колебательный повторяющийся процесс. В этом случае  $\omega_0$  совпадает с угловой скоростью вращения. Используют также и **частоту** колебаний:

$$\nu_0 = \frac{1}{T} = \frac{\omega_0}{2\pi}. \quad (24.26)$$

Частота  $\nu_0$  — это число колебаний, которое совершает система в единицу времени (в СИ — за секунду). Соответственно круговая частота равна числу колебаний за  $2\pi$  секунд. В СИ единицей частоты (круговой частоты) являются обратные секунды, для которых введено специальное название — герц ( $1 \text{ Гц} = 1/\text{с}$ ).

Итак, если система в момент времени  $t$  характеризуется значением переменной  $x(t)$ , то же самое значение переменная будет иметь через промежуток времени  $T = 2\pi/\omega_0$ :  $x(t) = x(t+T)$ . Это же значение, естественно, повторится через время  $2T, 3T$  и т. д.

В общем решении входят две произвольные постоянные ( $C_1, C_2$  или  $A, \alpha$ , значения которых должны определяться двумя **начальными условиями**. Обычно (хотя и необязательно) их роль играют начальные значения переменной  $x(0)$  и ее производной  $\dot{x}(0)$ .

Приведем пример. Пусть решение уравнения гармонических колебаний (24.19) описывает пружинный маятник. Значения произвольных постоянных зависят от способа, каким мы вывели маятник из состояния равновесия. Например, мы оттянули пружину на расстояние  $x_0$  и отпустили шарик без начальной скорости. В этом случае  $x(0) = x_0, \dot{x}(0) = 0$ . Подставляя  $t = 0$  в (24.19), находим:  $x(0) = x_0 = C_2$ . Решение, таким образом, имеет вид

$$x(t) = C_1 \sin \omega_0 t + x_0 \cos \omega_0 t.$$

Скорость груза находим дифференцированием:

$$\dot{x}(t) = C_1 \omega_0 \cos \omega_0 t - x_0 \omega_0 \sin \omega_0 t.$$

Подставляя сюда  $t = 0$ , находим вторую постоянную:  $\dot{x}(0) = C_1 \omega_0 = 0 \rightarrow C_1 = 0$ . Окончательно

$$x(t) = x_0 \cos \omega_0 t.$$

Сравнивая это уравнение с (24.23), находим, что  $x_0$  — это амплитуда колебаний, а его начальная фаза равна нулю:  $\alpha = 0$ .

Выведем теперь маятник из равновесия другим способом. Ударим по грузу, так что он приобретет начальную скорость  $v_0$ . Имеем тогда другие начальные условия:  $x(0) = 0, \dot{x}(0) = v_0$ . Так как  $x(0) = C_2 = 0$ , то наше решение имеет вид  $x(t) = C_1 \sin \omega_0 t$ . Скорость груза будет изменяться по закону  $\dot{x}(t) = C_1 \omega_0 \cos \omega_0 t$ . Подставим сюда  $t = 0$ :  $\dot{x}(0) = v_0 = C_1 \omega_0$ ,  $C_1 = v_0 / \omega_0$ . Окончательно получаем

$$x(t) = \frac{v_0}{\omega_0} \sin \omega_0 t = \frac{v_0}{\omega_0} \cos \left( \omega_0 t - \frac{\pi}{2} \right),$$

так что амплитуда колебаний равна  $A = v_0 / \omega_0$ , а начальная фаза  $\alpha = -\pi/2$ .

В общем случае, когда маятник отводится на расстояние  $x_0$  и отпускается с начальной скоростью  $v_0$ , связь начальных условий с амплитудой колебаний и начальной фазой имеет вид

$$A = \sqrt{x_0^2 + \frac{v_0^2}{\omega_0^2}} \quad \text{и} \quad \text{tg } \alpha = -\frac{v_0}{x_0 \omega_0}. \quad (24.27)$$

Дифференцируя решение (24.23) по времени, находим зависимость от времени, скорости и ускорения маятника:

$$\begin{cases} v(t) = \dot{x} = -A\omega_0 \sin(\omega_0 t + \alpha) = A\omega_0 \cos\left(\omega_0 t + \alpha + \frac{\pi}{2}\right), \\ a(t) = \ddot{x} = -A\omega_0^2 \cos(\omega_0 t + \alpha) = A\omega_0^2 \cos(\omega_0 t + \alpha + \pi). \end{cases} \quad (24.28)$$

Видно, что скорость и ускорение также изменяются по гармоническому закону, причем амплитуда скорости равна  $A\omega_0$ , а амплитуда ускорения —  $A\omega_0^2$ . Скорость опережает смещение по фазе на  $\pi/2$ , ускорение находится в противофазе по отношению к смещению. Это означает, что в тот момент, когда смещение достигает наибольшего значения, ускорение достигает наибольшего по величине отрицательного значения и наоборот.

**Задача 24.1.** В очень грубом приближении Землю можно представить себе в виде твердого шара радиусом  $R_{\oplus} = 6,38 \cdot 10^6$  м и массой  $M_{\oplus} = 6,0 \cdot 10^{24}$  кг. Пусть в нем пробурили узкую сквозную радиальную скважину, проходящую через центр Земли, и уронили в нее камешек без начальной скорости. При каком сферически симметричном распределении плотности Земли  $\rho(r)$  камешек будет совершать гармонические колебания? Найдите для этого случая период колебаний и амплитуду скорости камешка. Сопротивлением воздуха и вращением Земли пренебречь.

**Решение.** Всемирный закон тяготения Ньютона аналогичен закону Кулона в электростатике. Поэтому на камешек, находящийся на расстоянии  $r$  от центра Земли, будет действовать сила притяжения только той части массы Земли, которая находится внутри сферы радиусом  $r$ . Эта масса определяется соотношением

$$M(r) = 4\pi \int_0^r \rho(r) r^2 dr, \quad M(R_{\oplus}) = M_{\oplus}. \quad (24.29)$$

Уравнение движения камешка имеет вид

$$m\ddot{r} = -G \frac{M(r)m}{r^2}. \quad (24.30)$$

Уравнение гармонических колебаний с частотой  $\omega_{\oplus}$  (24.18) получается отсюда при линейной зависимости правой части (24.30) от расстояния до центра Земли:

$$G \frac{M(r)}{r^2} = \omega_{\oplus}^2 r. \quad (24.31)$$

Умножая обе части (24.31) на  $r^2$  и используя (24.29), приходим к соотношению

$$4\pi G \int_0^r \rho(r)r^2 dr = \omega_{\text{З}}^2 r^3. \quad (24.32)$$

Дифференцируя (24.32) по  $r$ , получаем, что плотность Земли должна быть постоянной:

$$\rho(r) = \rho_{\text{З}} = \frac{3}{4\pi G} \omega_{\text{З}}^2. \quad (24.33)$$

Постоянная плотность связана с массой Земли и ее радиусом соотношением  $M_{\text{З}} = \rho_{\text{З}}(4\pi R_{\text{З}}^3/3)$ . Отсюда получаем

$$\omega_{\text{З}} = \sqrt{\frac{GM_{\text{З}}}{R_{\text{З}}^3}} = \sqrt{\frac{g}{R_{\text{З}}}}, \quad (24.34)$$

где  $g = GM_{\text{З}}/R_{\text{З}}^2$  — ускорение свободного падения на поверхности Земли. Теперь находим период колебаний камешка, упавшего в скважину:

$$T_{\text{З}} = \frac{2\pi}{\omega_{\text{З}}} = 2\pi \sqrt{\frac{R_{\text{З}}}{g}}. \quad (24.35)$$

Подставляем сюда численные данные:

$$T_{\text{З}} = 2\pi \sqrt{\frac{6,38 \cdot 10^6}{9,81}} = 0,81 \cdot 10^3 \text{ с} \approx 1 \text{ ч } 24 \text{ мин.} \quad (24.36)$$

Амплитуда колебаний камешка равна радиусу Земли, значит, амплитуда скорости

$$v_{\text{З}} = \omega_{\text{З}} R_{\text{З}} = \sqrt{g R_{\text{З}}}. \quad (24.37)$$

Такую скорость набирает камешек к моменту прохождения центра Земли. Мы видим, что эта скорость совпадает с первой космической скоростью  $v_1$  (см. т. 1, гл. 6, разд. 6.6), численное значение которой  $v_1 = 7,9$  км/с. Таким образом, период  $T_1$  вращения спутника, летящего вблизи поверхности Земли, равен периоду колебаний камешка, падающего в скважине, проходящей через центр Земли:

$$T_1 = \frac{2\pi R_{\text{З}}}{v_1} = \frac{2\pi}{\omega_{\text{З}}} = T_{\text{З}}. \quad (24.38)$$

Конечно, это равенство — следствие предположения об однородности плотности нашей планеты, что заведомо не так. При другом законе зависимости плотности  $\rho(r)$  от расстояния колебания падающего камешка не будут гармоническими и их период будет отличаться от периода обращения спутника вокруг Земли. ■

## 24.3 Сохранение энергии при гармонических колебаниях

Умножим уравнение гармонических колебаний (24.18) на скорость изменения переменной  $x$ :

$$\dot{x} \ddot{x} + \omega_0^2 \dot{x} x = 0. \quad (24.39)$$

Каждое из слагаемых можно представить как производную:

$$\dot{x} \ddot{x} = \frac{d}{dt} \frac{\dot{x}^2}{2}, \quad \dot{x} x = \frac{d}{dt} \frac{x^2}{2}.$$

С учетом этого уравнения (24.39) можно записать в виде

$$\frac{d}{dt} \left( \frac{\dot{x}^2}{2} + \omega_0^2 \frac{x^2}{2} \right) = 0. \quad (24.40)$$

Отсюда следует, что величина в скобках не зависит от времени, т. е. сохраняется в процессе колебаний:

$$\frac{\dot{x}^2}{2} + \omega_0^2 \frac{x^2}{2} = \text{const}. \quad (24.41)$$

Для выяснения физического смысла сохраняющейся величины применим эти соотношения к пружинному маятнику, когда  $\omega_0^2 = k/m$ . Видим, что уравнение (24.41) можно записать в виде суммы кинетической энергии груза и потенциальной энергии сжатой пружины:

$$\boxed{\frac{m\dot{x}^2}{2} + \frac{kx^2}{2} = \text{const}.} \quad (24.42)$$

Таким образом, найденный закон сохранения есть не что иное, как **закон сохранения полной энергии колебательной системы**. Аналогично для электромагнитного контура переменная  $x = q$ ,  $\dot{x} = \dot{q} = I$  и  $\omega_0^2 = 1/LC$ . В этом случае соотношение (24.41) принимает вид

$$\boxed{\frac{LI^2}{2} + \frac{q^2}{2C} = \text{const}.} \quad (24.43)$$

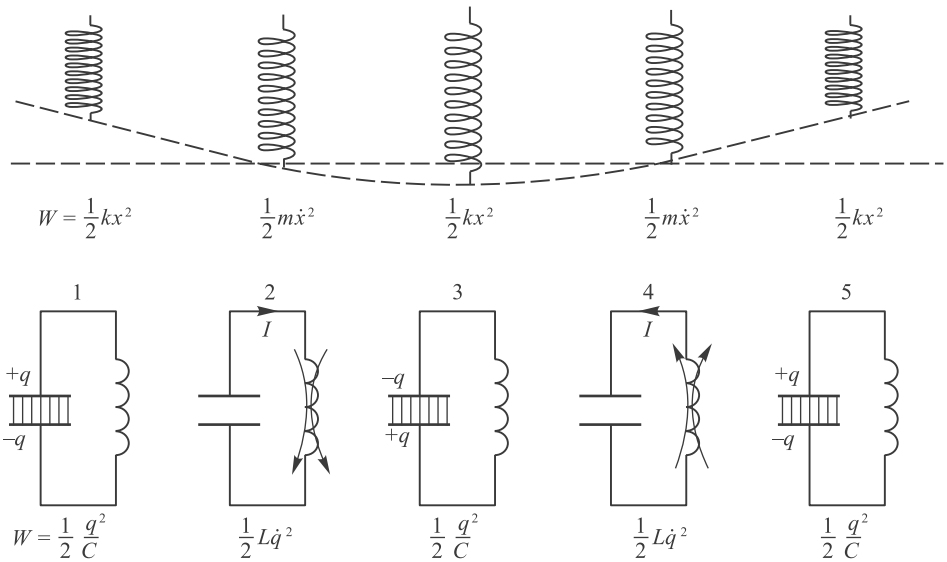
Первый член — энергия магнитного поля в катушке, второй — энергия электрического поля в конденсаторе. Снова мы получили, что сохраняется полная энергия в системе.

Возвращаясь к общей форме закона сохранения энергии (24.41) и подставляя сюда общее решение (24.23), получаем **законы изменения во времени кинетической и потенциальной энергий** (или их аналогов) и выражение для **сохраняющейся полной энергии колебаний**:

$$\boxed{\begin{aligned} T &= \frac{\dot{x}^2}{2} = \frac{A^2\omega_0^2}{2} \sin^2(\omega_0 t + \alpha), \\ U &= \frac{\omega_0^2 x^2}{2} = \frac{A^2\omega_0^2}{2} \cos^2(\omega_0 t + \alpha), \\ W &= T + U = \frac{A^2\omega_0^2}{2}. \end{aligned}} \quad (24.44)$$

Отсюда следует, что:

- кинетическая и потенциальная энергии — периодические функции времени с периодом, равным половине периода колебаний;
- кинетическая и потенциальная энергии колеблются в противофазе — когда кинетическая энергия достигает максимума, значение потенциальной энергии минимально и наоборот;
- в колебательной системе энергия периодически «перекачивается» из одной формы в другую, а полная энергия  $W = T + U$  сохраняется;



**Рис. 24.8.** Различные стадии изменения во времени кинетической и потенциальной энергий в колебательной системе

- полная энергия колебаний пропорциональна квадрату их амплитуды и квадрату частоты.

Сказанное проиллюстрировано на рис. 24.8, на котором показаны различные стадии изменения кинетической и потенциальной энергий для пружинного маятника и электромагнитного контура.

## 24.4 Сложение однонаправленных колебаний

Может случиться так, что осциллятор принимает участие в двух одинаково направленных колебаниях с разными амплитудами, частотами и начальными фазами. Рассмотрим сложение таких колебаний.

### Сложение колебаний с одинаковыми частотами

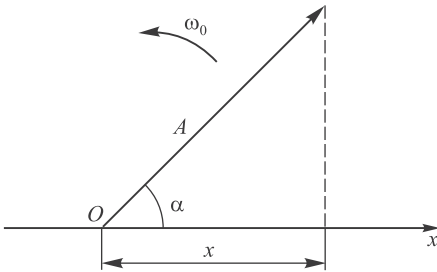
Для простоты рассмотрим сначала случай, когда частоты складываемых колебаний одинаковы. Общие решения складываемых гармонических колебаний имеют вид

$$\begin{aligned} x_1 &= A_1 \cos(\omega_0 t + \alpha_1), \\ x_2 &= A_2 \cos(\omega_0 t + \alpha_2), \end{aligned} \tag{24.45}$$

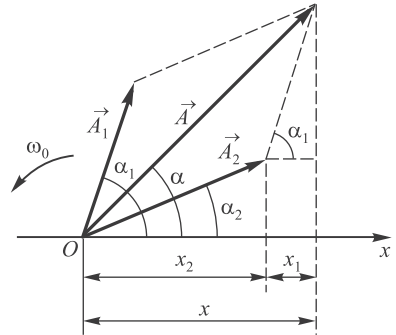
где  $x_1, x_2$  — переменные, описывающие колебания,  $A_1, A_2$  — их амплитуды, а  $\alpha_1, \alpha_2$  — начальные фазы. Результирующее колебание  $x = x_1 + x_2$  удобно найти с помощью **векторной диаграммы**. В этом методе используется аналогия между вращением и колебательным процессом.

Возьмем общее решение (24.23) для гармонического колебания. Выберем ось  $Ox$ . Из точки  $O$  отложим вектор длиной  $A$ , образующий с осью  $Ox$  угол  $\alpha$ . Если





**Рис. 24.9.** Векторная диаграмма гармонического колебания



**Рис. 24.10.** Векторная диаграмма для сложения одинаково направленных колебаний одинаковой частоты

привести этот вектор во вращение с угловой скоростью  $\omega_0$ , то проекция конца этого вектора будет перемещаться по оси  $Ox$  от  $+A$  до  $-A$ , причем величина проекции будет изменяться по закону

$$x(t) = A \cos(\omega_0 t + \alpha). \quad (24.46)$$

Таким образом, проекция конца вектора на ось  $Ox$  будет совершать гармонические колебания с амплитудой, равной длине вектора, с круговой частотой, равной угловой скорости вращения вектора, и с начальной фазой, равной углу, образуемому вектором с осью в начальный момент времени (рис. 24.9).

Применим теперь эту технику к сложению колебаний (24.45). Представим оба колебания векторами  $\vec{A}_1$  и  $\vec{A}_2$ . Возьмем их векторную сумму  $\vec{A} = \vec{A}_1 + \vec{A}_2$  (рис. 24.10). Проекция  $\vec{A}$  на ось  $Ox$  равна сумме проекций  $x = x_1 + x_2$ . Таким образом, вектор  $\vec{A}$  представляет собой результирующее колебание. Этот вектор вращается с той же угловой скоростью  $\omega_0$ , так что результирующее движение будет гармоническим колебанием с частотой  $\omega_0$ , амплитудой  $A$  и начальной фазой  $\alpha$ . Имеем согласно теореме косинусов

$$\begin{aligned} A^2 &= A_1^2 + A_2^2 + 2A_1A_2 \cos(\alpha_1 - \alpha_2), \\ \operatorname{tg} \alpha &= \frac{A_1 \sin \alpha_1 + A_2 \sin \alpha_2}{A_1 \cos \alpha_1 + A_2 \cos \alpha_2}. \end{aligned} \quad (24.47)$$

В частности, если фазы складываемых колебаний равны или различаются на четное кратное  $\pi$ , т. е.  $\alpha_1 - \alpha_2 = 2\pi n$ , то амплитуда результирующего колебания равна сумме амплитуд  $A = A_1 + A_2$ . Если складываемые колебания находятся в противофазе ( $\alpha_1 - \alpha_2 = 2\pi n + \pi$ ), то  $A = |A_1 - A_2|$ .

## Биения

В этом разделе мы рассмотрим случай сложения одинаково направленных гармонических колебаний с разными частотами. На практике особый интерес представляет случай, когда складываемые колебания мало различаются по частоте. Как мы увидим, в результате сложения этих колебаний получаются колебания с

периодически изменяющейся амплитудой. Периодические изменения амплитуды колебаний, возникающие при сложении двух гармонических колебаний с близкими частотами, называются *биениями*.

Для простоты рассмотрим случай, когда амплитуды складываемых колебаний равны  $A$ , а начальные фазы обоих колебаний равны нулю. Частоты складываемых колебаний равны соответственно  $\omega_1$  и  $\omega_2$ . Итак,

$$x_1 = A \cos \omega_1 t, \quad x_2 = A \cos \omega_2 t. \quad (24.48)$$

Складываем эти выражения и учитываем известную формулу тригонометрии:

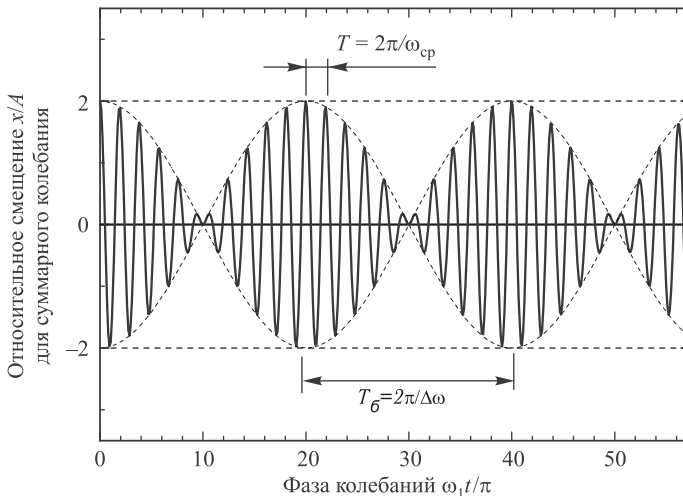
$$\begin{aligned} x &= A (\cos \omega_1 t + \cos \omega_2 t) = \\ &= 2A \cos \left( \frac{\omega_1 - \omega_2}{2} t \right) \cos \left( \frac{\omega_1 + \omega_2}{2} t \right). \end{aligned} \quad (24.49)$$

Второй множитель в правой части (24.49) описывает гармоническое колебание со средней частотой  $\omega_{\text{ср}} = (\omega_1 + \omega_2)/2$ . Если частоты  $\omega_1$  и  $\omega_2$  близки, т. е. разность частот мала ( $\Delta\omega = |\omega_1 - \omega_2| \ll \omega_{\text{ср}}$ ), то первый множитель меняется медленно по сравнению со вторым. Поэтому результирующее колебание

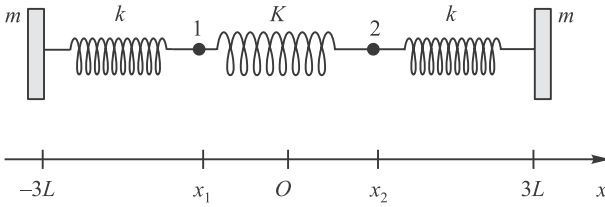
$$x = \left( 2A \cos \frac{\Delta\omega}{2} t \right) \cos \omega_{\text{ср}} t \quad (24.50)$$

можно рассматривать как гармоническое колебание с частотой  $\omega_{\text{ср}}$ , эффективная амплитуда  $A_{\text{эф}}$  которого изменяется со временем (рис. 24.11):

$$A_{\text{эф}} = \left| 2A \cos \frac{\Delta\omega}{2} t \right|. \quad (24.51)$$



**Рис. 24.11.** Биения при сложении колебаний с близкими частотами:  $\omega_2 = 1,1 \omega_1$ . Тонкие пунктирные кривые показывают зависимость от времени эффективной амплитуды колебаний  $\pm A_{\text{эф}}(t) = \pm 2A \cos(\Delta\omega t/2)$



**Рис. 24.12.** Пример связанных осцилляторов (колебания происходят вдоль оси  $Ox$ , сила тяжести не учитывается)

Частота пульсаций амплитуды (ее называют **частотой биений**) равна разности частот складываемых колебаний. Период биений равен:

$$T_6 = \frac{2\pi}{\Delta\omega}. \quad (24.52)$$

### Колебания двух связанных осцилляторов

Приведем поучительный пример системы, в которой возникают биения. Рассмотрим два груза массой  $m$  каждый, которые могут колебаться под действием двух одинаковых пружин с коэффициентами жесткости  $k$ . Пусть грузы соединены также мягкой пружиной с коэффициентом жесткости  $K \ll k$ . Будем полагать длины всех пружин в нерастянутом состоянии одинаковыми и равными  $2L$  (рис. 24.12). Тогда в положении равновесия координаты грузов равны  $x_{1,\text{рав}} = -L$ ,  $x_{2,\text{рав}} = L$ . При колебаниях координаты равны соответственно  $x_1(t)$ ,  $x_2(t)$ . Удлинения пружин можно записать как  $\Delta x_1 = x_1 + L$  (для левой),  $\Delta x_2 = x_2 - L$  (для правой) и  $\Delta x_3 = x_2 - x_1 - 2L$  (для средней).

Мы имеем дело с системой с двумя степенями свободы. Составим уравнения движения. На первый груз действуют:

- 1) сила  $F_1 = -k(x_1 + L)$  со стороны пружины  $k$ ;
- 2) сила  $F_{12} = K(x_2 - x_1 - 2L)$  со стороны пружины  $K$ .

Аналогичные силы  $F_2 = -k(x_2 - L)$  и  $F_{21} = -K(x_2 - x_1 - 2L)$  действуют на второй груз. Соответственно уравнения движения имеют вид

$$\begin{aligned} m\ddot{x}_1 &= -k(x_1 + L) + K(x_2 - x_1 - 2L), \\ m\ddot{x}_2 &= -k(x_2 - L) - K(x_2 - x_1 - 2L). \end{aligned} \quad (24.53)$$

На первый взгляд эти уравнения не слишком похожи на уравнения гармонических колебаний, потому что на колебания  $x_1$  оказывают влияния колебания  $x_2$  и наоборот. Поэтому преобразуем уравнения к новым переменным, колебания которых были бы независимыми (такие переменные называют **нормальными модами**). Именно введем новые переменные  $\xi_1$  и  $\xi_2$ :

$$\begin{aligned} \xi_1 &= \frac{x_1 + x_2}{\sqrt{2}}, & \xi_2 &= \frac{x_2 - x_1 - 2L}{\sqrt{2}}; \\ x_1 &= \frac{\xi_1 - \xi_2}{\sqrt{2}} - L, & x_2 &= \frac{\xi_1 + \xi_2}{\sqrt{2}} + L. \end{aligned} \quad (24.54)$$

Как легко убедиться, система находится в равновесии при  $\xi_1 = \xi_2 = 0$ .

В этих переменных уравнения (24.53) принимают вид

$$\begin{aligned} m(\ddot{\xi}_1 - \ddot{\xi}_2) &= -k(\xi_1 - \xi_2) + 2K\xi_2, \\ m(\ddot{\xi}_1 + \ddot{\xi}_2) &= -k(\xi_1 + \xi_2) - 2K\xi_2. \end{aligned} \quad (24.55)$$

Складывая и вычитая эти уравнения, приходим к паре независимых уравнений для колебаний нормальных мод:

$$\begin{aligned} m\ddot{\xi}_1 &= -k\xi_1, \\ m\ddot{\xi}_2 &= -(k + 2K)\xi_2. \end{aligned} \quad (24.56)$$

При отсутствии соединительной пружины  $K$  оба маятника совершают независимые гармонические колебания с одинаковой частотой  $\omega_0 = \sqrt{k/m}$ . При включении взаимодействия маятников частоты нормальных колебаний расщепляются на две частоты  $\omega_1$  и  $\omega_2$ . В данной задаче

$$\omega_1 = \omega_0 = \sqrt{\frac{k}{m}}, \quad \omega_2 = \sqrt{\frac{k + 2K}{m}}. \quad (24.57)$$

Для величины расщепления получаем уравнение

$$\Delta\omega = \omega_2 - \omega_1 = \frac{\omega_2^2 - \omega_1^2}{\omega_1 + \omega_2} = \frac{K}{m\omega_{cp}}, \quad (24.58)$$

где  $\omega_{cp} = (\omega_1 + \omega_2)/2$ . При  $K \ll \omega_0$  имеем  $\omega_{cp} \approx \omega_0$  и

$$\Delta\omega \approx \frac{K}{k} \omega_0. \quad (24.59)$$

Соответственно мы получаем общие решения для нормальных мод:

$$\xi_1 = A_1 \cos(\omega_1 t + \alpha_1), \quad \xi_2 = A_2 \cos(\omega_2 t + \alpha_2). \quad (24.60)$$

Общие решения для координат колеблющихся точек следуют из (24.54) и (24.60):

$$\begin{aligned} x_1 &= A_1 \cos(\omega_1 t + \alpha_1) - A_2 \cos(\omega_2 t + \alpha_2) - L, \\ x_2 &= A_1 \cos(\omega_1 t + \alpha_1) + A_2 \cos(\omega_2 t + \alpha_2) + L. \end{aligned} \quad (24.61)$$

Для примера рассмотрим случай, когда в начальный момент  $t = 0$  вторая масса оттягивается на расстояние  $a$  от положения равновесия и отпускается с нулевой начальной скоростью, а первая масса остается в положении равновесия:

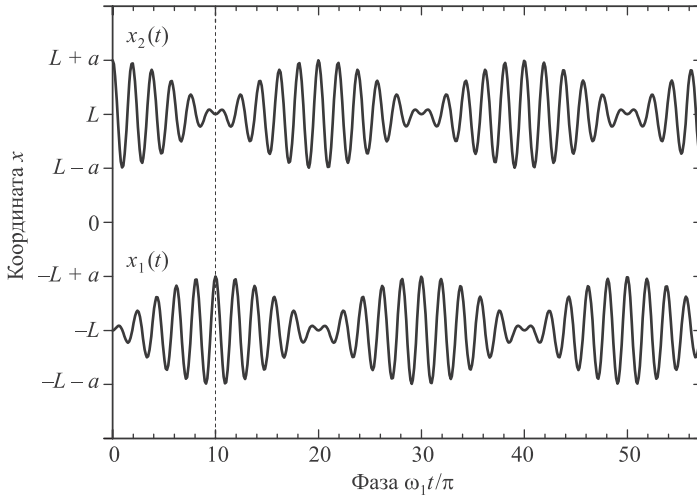
$$\begin{aligned} x_1(0) &= -L, & \dot{x}_1(0) &= 0, \\ x_2(0) &= L + a, & \dot{x}_2(0) &= 0. \end{aligned} \quad (24.62)$$

Этому соответствуют начальные условия для нормальных мод:

$$\xi_1(0) = \xi_2(0) = \frac{a}{2}, \quad \dot{\xi}_1(0) = \dot{\xi}_2(0) = 0. \quad (24.63)$$

Уравнения колебаний с такими начальными условиями мы уже решали и можем сразу написать ответ:

$$\xi_1 = \frac{a}{2} \cos \omega_1 t, \quad \xi_2 = \frac{a}{2} \cos \omega_2 t. \quad (24.64)$$



**Рис. 24.13.** Биения в системе двух связанных осцилляторов при  $\omega_2 = 1,1\omega_1$

Таким образом, мы нашли амплитуды  $A_1 = A_2 = a/2$  и фазы  $\alpha_1 = \alpha_2 = 0$  нормальных колебаний (см. (24.60)). Подставляя их в (24.61), получаем решения, описывающие биения наших масс около их положений равновесия  $\pm L$ :

$$\begin{aligned} x_1 &= -L + \frac{a}{2} (\cos \omega_1 t - \cos \omega_2 t) = -L + a \sin(\Delta\omega t/2) \sin \omega_{cp} t, \\ x_2 &= L + \frac{a}{2} (\cos \omega_1 t + \cos \omega_2 t) = L + a \cos(\Delta\omega t/2) \cos \omega_{cp} t. \end{aligned} \quad (24.65)$$

Графики функций  $x_1(t)$ ,  $x_2(t)$  показаны на рис. 24.13. В начальный момент времени смещается лишь второй груз. Затем начинает колебаться первый, а амплитуда колебаний второго уменьшается. Через время  $t = \pi/\Delta\omega$  второй груз останавливается, а колебания первого происходят с максимально возможной амплитудой (этот момент показан на рис. 24.13 вертикальной пунктирной линией). Произошла «перекачка» энергии от второго маятника к первому. Затем процесс «перекачки» энергии идет в обратном направлении и к моменту  $t = 2\pi/\Delta\omega$  второй маятник колеблется с максимальной амплитудой, а первый покоится.

Выясним теперь физический смысл нормальных мод, каждая из которых соответствует независимому гармоническому колебанию. Если возбуждены колебания только первой из них ( $\xi_1$ ), то  $A_2 = 0$  и, как следует из общего решения (24.61),

$$\begin{aligned} x_1 &= -L + A_1 \cos(\omega_1 t + \alpha_1), \\ x_2 &= L + A_1 \cos(\omega_1 t + \alpha_1). \end{aligned} \quad (24.66)$$

В этом случае грузы смещаются из положения равновесия в одну сторону на одинаковые расстояния. Скорости их также одинаковы по величине и направлению. Соединительная пружина колеблется вместе с грузами, но остается нерастянутой и потому не оказывает влияния, так что частота колебаний совпадает с частотой колебаний несвязанных маятников.

Возбуждение только второй ( $\xi_2$ ) нормальной моды означает, что  $A_1 = 0$ :

$$\begin{aligned}x_1 &= -L - A_2 \cos(\omega_2 t + \alpha_2), \\x_2 &= L + A_2 \cos(\omega_2 t + \alpha_2).\end{aligned}\tag{24.67}$$

Эта нормальная мода соответствует такому колебанию, когда оба груза смещаются на одинаковые расстояния от своих положений равновесия, но в противоположные стороны. Скорости движения грузов также равны по величине и противоположны по направлению, так что центр масс грузов остается неподвижным. Колебания происходят под действием пружин с жесткостью  $k$ , к которым добавляется соединительная пружина с жесткостью  $K$ . Как следствие, частота таких колебаний больше частоты колебаний несвязанных осцилляторов.

В разобранным случае мы познакомились с нормальными модами и выяснили, что их частоты сдвигаются по сравнению с частотами колебаний несвязанных маятников. Любое другое колебательное движение системы можно представить как суперпозицию нормальных мод. Аналогичным образом можно рассмотреть цепочку из множества связанных друг с другом осцилляторов и изучить их нормальные колебания. Такая система представляет собой модель кристаллической решетки.

## 24.5 Сложение взаимно перпендикулярных колебаний

В этом разделе мы рассмотрим сложение двух гармонических колебаний одинаковой частоты  $\omega$ , происходящих во взаимно перпендикулярных направлениях вдоль осей  $x$  и  $y$ . Выберем начало отсчета времени так, чтобы начальная фаза первого колебания была равна нулю:

$$x = A \cos \omega t, \quad y = B \cos(\omega t + \alpha),\tag{24.68}$$

где  $\alpha$  — разность фаз этих колебаний. Найдем уравнение траектории материальной точки, участвующей в этих взаимно перпендикулярных колебаниях, т. е. функцию  $y = y(x)$ .

Рассмотрим сначала частные случаи. Пусть разность фаз равна нулю:  $\alpha = 0$ . Тогда

$$y = \frac{B}{A} x,\tag{24.69}$$

т. е. траектория точки представляет собой прямую. Аналогичная прямая, но проходящая в двух других квадрантах, получится при колебаниях с разностью фаз  $\alpha = \pi$ :

$$y = B \cos(\omega t + \pi) = -B \cos(\omega t) = -\frac{B}{A} x.\tag{24.70}$$

При разности фаз  $\alpha = \pi/2$  находим:

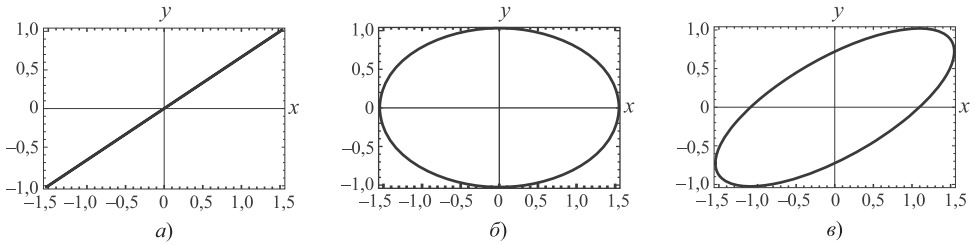
$$y = B \cos\left(\omega t + \frac{\pi}{2}\right) = -B \sin \omega t.\tag{24.71}$$

Из (24.68) и (24.71) следует уравнение эллипса

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1,\tag{24.72}$$

причем точка движется по часовой стрелке. При разности фаз  $\alpha = -\pi/2$  находим:

$$y = B \cos\left(\omega t - \frac{\pi}{2}\right) = B \sin \omega t.\tag{24.73}$$



**Рис. 24.14.** Траектории материальной точки, колеблющейся с одинаковыми частотами и отношением амплитуд  $A/B = 1,5$  во взаимно перпендикулярных направлениях при различных разностях фаз: а)  $\alpha_1 = \alpha_2 = 0$ , б)  $\alpha_1 = 0, \alpha_2 = \pi/2$ , в)  $\alpha_1 = 0, \alpha_2 = \pi/4$

Это приводит к той же эллиптической траектории (см. (24.72)), только вращение в этом случае происходит против часовой стрелки.

В общем случае произвольной разности фаз  $\alpha$  траектория также будет представлять собой эллипс, но с повернутыми осями (рис. 24.14).

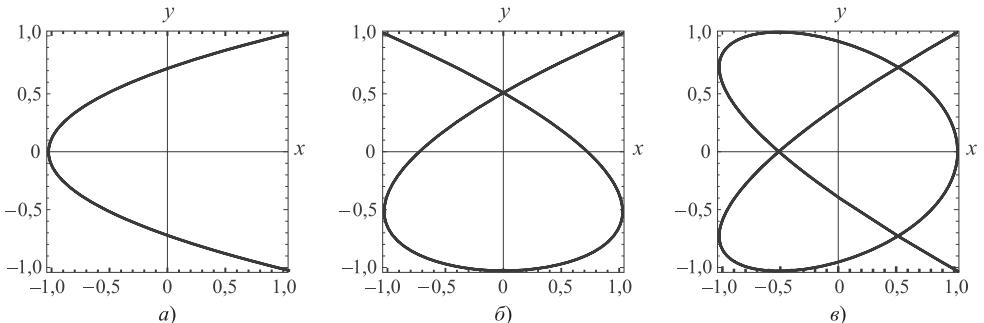
Если частоты взаимно перпендикулярных колебаний не одинаковы, то траектория результирующего движения имеет довольно сложный вид. Замкнутые траектории, описываемые точкой, совершающей одновременно два взаимно перпендикулярных колебания, называются фигурами Лиссажу.

Одна из простейших фигур Лиссажу получается при отношении частот 2:1 и нулевых начальных фазах обоих колебаний  $\alpha_1 = \alpha_2 = 0$ :

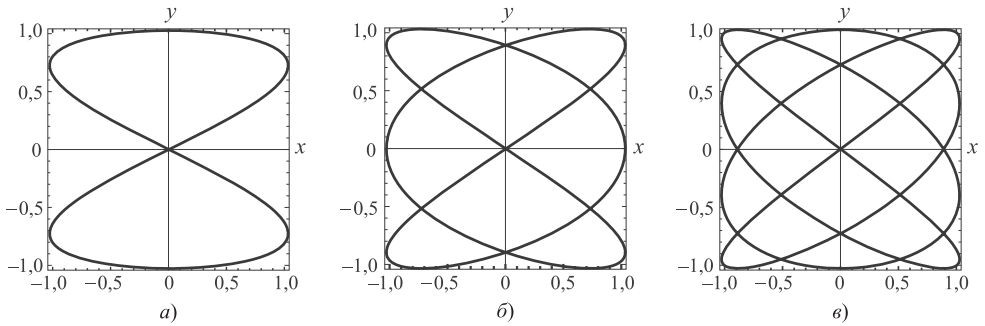
$$x = A \cos(2\omega t), \quad y = B \cos \omega t, \quad \text{откуда}$$

$$x = A (2 \cos^2 \omega t - 1) = \frac{2A}{B^2} y^2 - A. \quad (24.74)$$

Мы получили уравнение параболы. Форма фигур Лиссажу зависит от отношения частот складываемых колебаний и разности фаз между ними. Примеры показаны на рис. 24.15 и 24.16.



**Рис. 24.15.** Примеры фигур Лиссажу для разных отношений частот и для нулевых значений начальных фаз каждого колебания: а)  $\omega_1 : \omega_2 = 2 : 1$ ; б)  $\omega_1 : \omega_2 = 3 : 2$ ; в)  $\omega_1 : \omega_2 = 4 : 3$ . Координата  $x$  выражена в единицах амплитуды колебаний  $A$ , координата  $y$  — в единицах амплитуды  $B$ . Частота  $\omega_1$  относится к колебаниям вдоль оси  $x$ , частота  $\omega_2$  — вдоль оси  $y$



**Рис. 24.16.** Примеры фигур Лиссажу для тех же отношений частот, что и на рис. 24.15, и при одинаковых, но ненулевых начальных фазах каждого колебания:  $\alpha_1 = \alpha_2 = -\pi/2$ , т. е. при  $x = A \sin \omega_1 t, y = B \sin \omega_2 t$

## 24.6 Свободные затухающие колебания

Гармонические колебания, существующие вечно, являются одной из физических абстракций. В реальных системах колебания по прошествии некоторого времени затухают из-за диссипации энергии. Таким образом, представлением о гармонических колебаниях можно пользоваться лишь для времен, малых по сравнению с характерным временем затухания. Затухание колебаний всегда будет наблюдаться в системах с трением.

### Уравнение затухающих колебаний

Рассмотрим в качестве примера пружинный маятник, помещенный в вязкую среду. Помимо силы упругости на тело будет действовать сила сопротивления, пропорциональная скорости  $F_r = -r\dot{x}$ , где  $r$  — коэффициент, зависящий от вязкости среды, размеров и формы тела. Поэтому уравнение движения примет вид

$$m\ddot{x} = -kx - r\dot{x}, \tag{24.75}$$

или

$$\ddot{x} + 2\beta\dot{x} + \omega_0^2 x = 0, \quad \omega_0 = \sqrt{\frac{k}{m}}, \quad \beta = \frac{r}{2m}. \tag{24.76}$$

Другой пример — электромагнитный контур. Если помимо конденсатора  $C$  и индуктивности  $L$  в контуре имеется еще и активное сопротивление  $R$ , то ЭДС самоиндукции равна сумме напряжения на конденсаторе и падения напряжения на сопротивлении. Поэтому вместо уравнений (24.15) получим

$$I = \frac{dq}{dt}, \quad -L \frac{dI}{dt} = \frac{q}{C} + IR. \tag{24.77}$$

Подставляем первое уравнение во второе:

$$L \frac{d^2q}{dt^2} + \frac{q}{C} + R \frac{dq}{dt} = 0, \tag{24.78}$$



или

$$\ddot{q} + 2\beta\dot{q} + \omega_0^2 q = 0, \quad \omega_0 = \frac{1}{\sqrt{LC}}, \quad \beta = \frac{R}{2L}. \quad (24.79)$$

Напомним, что комбинация  $L/R$  уже встречалась нам в теории электромагнетизма, где она характеризовала характерное время затухания (появления) экстратовков замыкания-размыкания. Таким образом, величина  $\beta$  имеет размерность  $[\beta] = \text{с}^{-1}$ , совпадающую с размерностью частоты.

## Анализ решений

Итак, в обоих рассмотренных случаях дифференциальное уравнение свободных затухающих колебаний линейной системы имеет вид

$$\ddot{x} + 2\beta\dot{x} + \omega_0^2 x = 0, \quad (24.80)$$

где  $x$  — переменный физический параметр (смещение, заряд, ток),  $\beta$  — коэффициент затухания, а  $\omega_0$  — циклическая частота свободных (собственных) незатухающих колебаний (при  $\beta = 0$ , т. е. при отсутствии потерь энергии). Сведем новую задачу к предыдущей. Для этого вместо переменной  $x$  определим новую переменную  $X$ , связанную с  $x$  соотношением

$$x(t) = e^{-\beta t} X(t). \quad (24.81)$$

Дифференцируем функцию  $x(t)$ :

$$\begin{aligned} \dot{x} &= X \frac{d}{dt} e^{-\beta t} + \dot{X} e^{-\beta t} = e^{-\beta t} (\dot{X} - \beta X), \\ \ddot{x} &= X \frac{d^2}{dt^2} e^{-\beta t} + 2\dot{X} \frac{d}{dt} e^{-\beta t} + \ddot{X} e^{-\beta t} = \\ &= e^{-\beta t} (\ddot{X} - 2\beta\dot{X} + \beta^2 X). \end{aligned} \quad (24.82)$$

Подставляем эти выражения в уравнение (24.80):

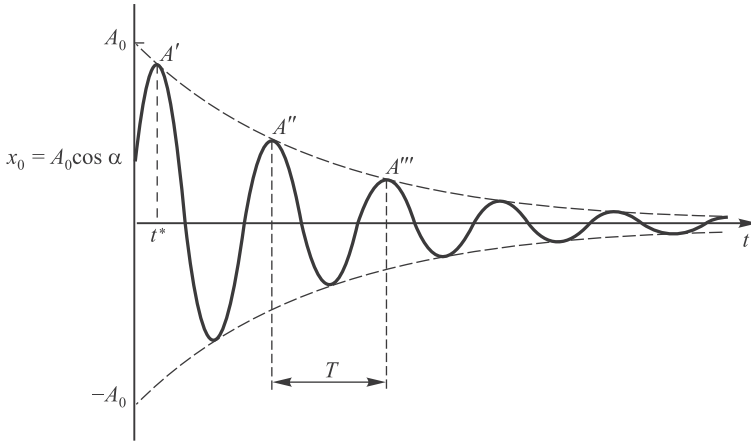
$$e^{-\beta t} [\ddot{X} - 2\beta\dot{X} + \beta^2 X + 2\beta(\dot{X} - \beta X) + \omega_0^2 X] = 0. \quad (24.83)$$

Выражение в квадратных скобках должно быть равно нулю. Замечаем, что в этом выражении сокращаются члены с первой производной  $\dot{X}$ . Получаем в итоге дифференциальное уравнение для функции  $X(t)$

$$\ddot{X} + [\omega_0^2 - \beta^2] X = 0. \quad (24.84)$$

Здесь возможны два случая. Пусть сначала  $\beta < \omega_0$ . Тогда можно ввести параметр  $\omega^2 = \omega_0^2 - \beta^2$ , так что уравнение (24.84) примет вид

$$\ddot{X} + \omega^2 X = 0.$$



**Рис. 24.17.** Свободные затухающие колебания при  $\beta = 0,1\omega_0$ ; пунктирные кривые соответствуют функции  $A(t) = A_0 \exp(-\beta t)$

Но это — стандартное уравнение гармонических колебаний, общее решение которого мы знаем:  $X = A_0 \cos(\omega t + \alpha)$ . Значит, мы нашли общее решение уравнения затухающих колебаний (24.80):

$$x(t) = A_0 e^{-\beta t} \cos(\omega t + \alpha), \quad \omega = \sqrt{\omega_0^2 - \beta^2}. \quad (24.85)$$

Во многих системах коэффициент затухания мал по сравнению с собственной частотой колебаний:  $\beta \ll \omega_0$ . Тогда движение системы можно рассматривать как почти гармоническое колебание с частотой  $\omega$  и с амплитудой, изменяющейся по закону  $A(t) = A_0 e^{-\beta t}$  (рис. 24.17). Коэффициент затухания  $\beta$  определяет скорость уменьшения амплитуды колебаний: он обратен по величине промежутку времени, за который амплитуда уменьшается в  $e$  раз.

Период затухающих колебаний равен:

$$T = \frac{2\pi}{\omega} = \frac{2\pi}{\sqrt{\omega_0^2 - \beta^2}}. \quad (24.86)$$

Пусть первое наибольшее отклонение  $A'$  достигается в момент времени  $t = t^*$ . Последующие наибольшие отклонения ( $A''$ ,  $A'''$  и т. д. — см. рис. 24.17) образуют геометрическую прогрессию:

$$A' = A_0 e^{-\beta t^*}, \quad A'' = A_0 e^{-\beta(t^*+T)}, \quad A''' = A_0 e^{-\beta(t^*+2T)}, \dots \quad (24.87)$$

Отношение значений амплитуд, соответствующих моментам времени, отличающимся на период колебаний  $T$ , равно:

$$\frac{A(t^*)}{A(t^* + T)} = e^{\beta T}. \quad (24.88)$$

Величина  $\exp(\beta T)$  называется **декрементом затухания**. Логарифм этого отношения называется **логарифмическим декрементом затухания**:

$$\lambda = \ln \frac{A(t^*)}{A(t^* + T)} = \ln(e^{\beta T}) = \beta T. \quad (24.89)$$

Определим количество колебаний, которые совершит система за **время релаксации**  $\tau = 1/\beta$ . За это время амплитуда уменьшается в  $e$  раз, а число колебаний равно:

$$N_e = \frac{\tau}{T} = \frac{1}{\beta T} = \frac{1}{\lambda}. \quad (24.90)$$

Следовательно, логарифмический декремент затухания обратен по величине числу колебаний, совершаемых за время релаксации.

Для характеристики колебательной системы часто используется величина, называемая **добротностью**:

$$Q = \frac{\pi}{\lambda} = \pi N_e. \quad (24.91)$$

Добротность пропорциональна числу колебаний  $N_e$ , совершаемых системой за время релаксации  $\tau$ . Например, для электромагнитного контура при  $\beta \ll \omega_0$  находим:

$$\begin{aligned} T &= \frac{2\pi}{\sqrt{\omega_0^2 - \beta^2}} \approx \frac{2\pi}{\omega_0} = 2\pi\sqrt{LC}, \\ \tau &= \frac{1}{\beta} = \frac{2L}{R}, \\ \lambda &= \beta T = \pi R \sqrt{\frac{C}{L}}, \\ Q &= \frac{\pi}{\lambda} = \frac{1}{R} \sqrt{\frac{L}{C}}. \end{aligned} \quad (24.92)$$

Мы видели, что полная энергия в колеблющейся системе пропорциональна квадрату амплитуды. При малом затухании ( $\beta \ll \omega_0$ ) имеем

$$E = E_0 e^{-2\beta t}, \quad (24.93)$$

где  $E_0$  — значение полной энергии колеблющейся системы в начальный момент времени. Можно определить изменение энергии за период  $T$ :

$$\Delta E \approx -2\beta T E. \quad (24.94)$$

Следовательно,

$$\frac{E}{(-\Delta E)} = \frac{Q}{2\pi}, \quad (24.95)$$

т.е. при слабом затухании добротность, с точностью до множителя  $1/(2\pi)$ , равна отношению полной энергии, запасенной в колебательной системе в данный момент времени, к убыли энергии за один период колебаний.

При увеличении затухания частота колебаний  $\omega = \sqrt{\omega_0^2 - \beta^2}$  уменьшается, а период колебаний растет — система может совершить всего одно колебание, прежде чем они окончательно затухнут. Такой случай показан на рис. 24.18 сплошной кривой, рассчитанной по уравнению (24.85) при коэффициенте затухания  $\beta = 0,5\omega_0$  для начальных условий  $x(0) = x_0$ ,  $\dot{x}(0) = 0$ . В предельном случае  $\beta \rightarrow \omega_0$  период обращается в бесконечность, т.е. движение перестает быть периодическим.

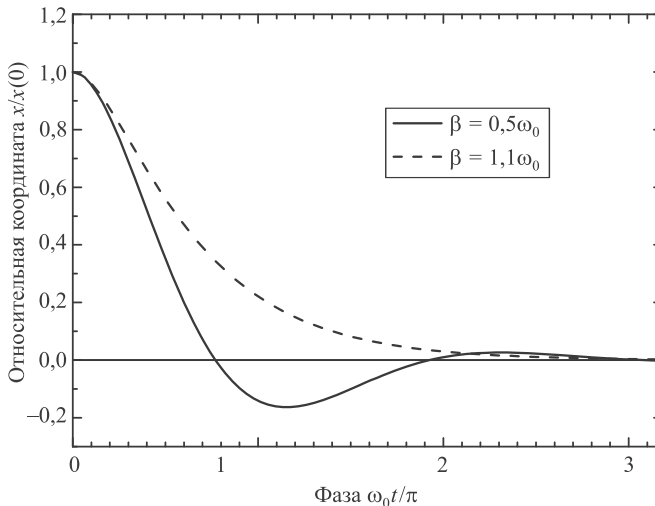
Если коэффициент затухания превышает частоту собственных колебаний ( $\beta > \omega_0$ ), то решение уравнения (24.84) выражается уже не через тригонометрические функции, описывающие колебания, а представляется в виде экспонент:

$$X(t) = C_1 \exp\left(-\sqrt{\beta^2 - \omega_0^2} t\right) + C_2 \exp\left(\sqrt{\beta^2 - \omega_0^2} t\right), \quad (24.96)$$

где коэффициенты  $C_1$  и  $C_2$  должны быть определены из начальных условий. При тех же условиях  $x(0) = x_0$ ,  $\dot{x}(0) = 0$  решение имеет вид

$$x(t) = x_0 e^{-\beta t} \left[ \operatorname{ch}\left(\sqrt{\beta^2 - \omega_0^2} t\right) + \frac{\beta}{\sqrt{\beta^2 - \omega_0^2}} \operatorname{sh}\left(\sqrt{\beta^2 - \omega_0^2} t\right) \right]. \quad (24.97)$$

Соответствующий этому решению график показан на рис. 24.18 штриховой кривой при коэффициенте затухания  $\beta = 1,1\omega_0$ . Видно, что при  $\beta > \omega_0$  движение носит аperiodический характер — выведенная из положения равновесия система возвращается в положение равновесия, не совершая колебаний.



**Рис. 24.18.** Затухание колебаний при  $\beta = 0,5\omega_0$  (сплошная линия) и  $\beta = 1,1\omega_0$  (пунктирная линия)

## 24.7 Вынужденные колебания

При *вынужденных колебаниях* система колеблется под действием периодической внешней (вынуждающей) силы. За счет работы этой силы компенсируются потери энергии системы. Частота вынужденных колебаний зависит от частоты изменения внешней силы (вынуждающей частоты). Практически наиболее интересным является случай, когда вынуждающая сила изменяется по гармоническому закону  $F_{\text{вын}} = F_m \cos \omega t$ . Зависимость амплитуды вынужденных колебаний от частоты вынуждающей силы приводит к тому, что при некоторой определенной для данной системы частоте амплитуда колебаний достигает максимального значения. Это явление называется **резонансом**. Явление резонанса используется для усиления колебаний, например электрических. Однако при конструировании машин и сооружений необходимо учитывать явление резонанса, чтобы предотвратить нежелательные последствия резонансного увеличения амплитуды вынужденных колебаний.

Для пружинного маятника уравнение вынужденного колебательного движения имеет вид

$$m\ddot{x} = -kx - r\dot{x} + F_m \cos \omega t, \quad (24.98)$$

или

$$\ddot{x} + 2\beta\dot{x} + \omega_0^2 x = a_m \cos \omega t, \quad (24.99)$$

где  $2\beta = r/m$ ;  $\omega_0^2 = k/m$ ;  $a_m = F_m/m$  и  $\omega$  — вынуждающая частота.

Если рассматривать электрический колебательный контур, то компенсировать потери энергии в контуре можно с помощью подводимой извне периодически изменяющейся по гармоническому закону ЭДС, или переменного напряжения:

$$U = U_m \cos \omega t. \quad (24.100)$$

Уравнение колебаний в контуре (рис. 24.19) запишем, используя закон Ома для замкнутой цепи:

$$IR = -\frac{q}{C} - L \frac{dI}{dt} + U_m \cos \omega t, \quad (24.101)$$

или, с учетом  $I = \dot{q}$ ,

$$\ddot{q} + 2\beta\dot{q} + \omega_0^2 q = a_m \cos \omega t, \quad (24.102)$$

где  $\omega_0^2 = 1/LC$  — собственная частота контура;  $\beta = R/2L$  — коэффициент затухания, а  $a_m = U_m/L$ .

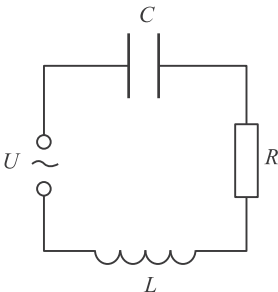


Рис. 24.19. Вынужденные колебания в электромагнитном контуре

Таким образом, вынужденные колебания в электрическом контуре описываются тем же самым линейным неоднородным дифференциальным уравнением второго порядка, что и колебания пружинного маятника в вязкой среде. Предположим, что нам известно хотя бы одно решение этого уравнения — некое частное решение  $q_{\text{вын}}(t)$ . Тогда разность любого другого решения  $q(t)$  и этого частного решения  $q_{\text{вын}}$  будет удовлетворять однородному уравнению (с нулем в правой части), которое мы подробно изучили в предыдущем разделе. Поэтому общее решение уравнения (24.102) может быть записано как

$$q(t) = q_{\text{вын}}(t) + q_0 e^{-\beta t} \cos(\omega_{\text{зат}} t + \alpha), \quad (24.103)$$

где  $\omega_{\text{зат}} = \sqrt{\omega_0^2 - \beta^2}$  — частота свободных затухающих колебаний. С течением времени из-за экспоненциального множителя  $e^{-\beta t}$  роль второго слагаемого уменьшается (оно важно на начальной стадии установления колебаний). По прошествии достаточно большого времени  $t \gg 1/\beta$  этим вторым слагаемым можно пренебречь, сохраняя лишь первое. Таким образом, задача сводится к нахождению хотя бы одного частного решения уравнения (24.102).

Частное решение неоднородного уравнения будем искать в виде гармонической функции, частота изменения которой совпадает с частотой вынуждающей силы:

$$q_{\text{вын}} = C_1 \cos \omega t + C_2 \sin \omega t. \quad (24.104)$$

Подставим  $q_{\text{вын}}$  в уравнение (24.102):

$$\begin{aligned} -\omega^2(C_1 \cos \omega t + C_2 \sin \omega t) + 2\beta\omega(-C_1 \sin \omega t + C_2 \cos \omega t) + \\ + \omega_0^2(C_1 \cos \omega t + C_2 \sin \omega t) - a_m \cos \omega t = 0. \end{aligned} \quad (24.105)$$

Так как функции синуса и косинуса линейно независимы, коэффициенты при них в левой части (24.105) должны быть равны нулю:

$$\begin{aligned} C_1(\omega_0^2 - \omega^2) + C_2 2\beta\omega - a_m = 0, \\ -C_1 2\beta\omega + C_2(\omega_0^2 - \omega^2) = 0. \end{aligned} \quad (24.106)$$

Решение этой системы имеет вид

$$\begin{aligned} C_1 &= a_m \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2}, \\ C_2 &= a_m \frac{2\beta\omega}{(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2}. \end{aligned} \quad (24.107)$$

Решение (24.104) с коэффициентами (24.107) записываем в стандартном виде:

$$\boxed{q_{\text{вын}} = A \cos(\omega t + \varphi)}, \quad (24.108)$$

где

$$\boxed{A = \sqrt{C_1^2 + C_2^2} = \frac{a_m}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2}}} \quad (24.109)$$

и

$$\boxed{\operatorname{tg}\varphi = -\frac{C_2}{C_1} = \frac{2\beta\omega}{\omega^2 - \omega_0^2}}. \quad (24.110)$$

Рассмотрим отклик системы на изменение частоты внешней силы. Под квадратным корнем в выражении для амплитуды стоит квадратичная функция частоты

$$f(\omega) = (\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2.$$

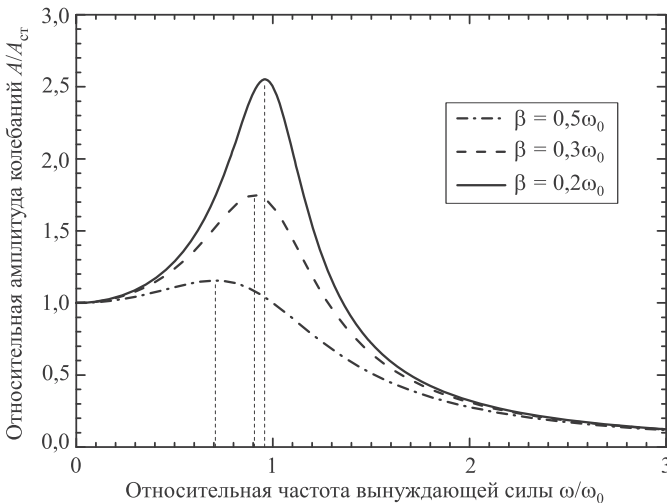
При некотором значении частоты  $\omega$  эта функция имеет минимум (а значит, амплитуда имеет максимум). Для нахождения точки минимума дифференцируем  $f(\omega)$  по  $\omega$  и приравниваем производную нулю. В итоге получаем для **резонансной частоты**

$$\boxed{\omega_{\text{рез}} = \sqrt{\omega_0^2 - 2\beta^2}}. \quad (24.111)$$

и

$$\boxed{A_{\text{рез}} = \frac{a_m}{2\beta\sqrt{\omega_0^2 - \beta^2}}}. \quad (24.112)$$

Следует отметить, что при  $\beta \ll \omega_0$  значение резонансной частоты  $\omega_{\text{рез}}$  практически совпадает с собственной частотой  $\omega_0$  колебательной системы. Поскольку  $\beta$  стоит в знаменателе выражения для  $A_{\text{рез}}$ , амплитуда колебаний при резонансе растет с уменьшением затухания. На графике рис. 24.20 видно, что чем меньше затухание, тем выше максимум резонансной кривой и тем правее он лежит.



**Рис. 24.20.** Зависимость относительной амплитуды  $A/A_{\text{ст}}$  вынужденных колебаний от относительной частоты  $\omega/\omega_0$  вынуждающей силы для разных значений коэффициента затухания  $\beta$ ; вертикальные пунктирные линии показывают значения резонансных частот

При увеличении частоты внешнего воздействия амплитуда стремится к нулю:

$$A \sim \frac{a_m}{\omega^2}, \quad \omega \gg \omega_0.$$

Физически это понятно: система обладает некой инерционностью и не успевает следовать за быстрыми изменениями внешнего воздействия. В другом предельном случае малой внешней частоты  $\omega \rightarrow 0$  мы имеем дело со статическим случаем — действием постоянной внешней силы  $F_m$  на пружинный маятник, подсоединением контура к источнику с постоянным напряжением  $U_m$ . В этом случае предельное значение амплитуды вынужденных колебаний равно:

$$A_{\text{ст}} = \frac{a_m}{\omega_0^2}$$

и не зависит от затухания. Последнее вполне естественно, так как затухание (сила сопротивления) проявляется только при движении системы, а не в статическом пределе. В случае механических колебаний

$$A_{\text{ст}} = \frac{F_m}{m\omega_0^2} = \frac{F_m}{k}, \quad (24.113)$$

что равно удлинению пружины под действием постоянной силы  $F_m$ . В случае электромагнитных колебаний в контуре

$$A_{\text{ст}} = \frac{U_m}{L\omega_0^2} = U_m C = q_m, \quad (24.114)$$

что равно заряду на конденсаторе при подсоединении его к источнику постоянного напряжения  $U_m$ .

Найдем отношение резонансной амплитуды к статической:

$$\frac{A_{\text{рез}}}{A_{\text{ст}}} = \frac{\omega_0^2}{2\beta\sqrt{\omega_0^2 - \beta^2}} \approx \frac{\omega_0}{2\beta} = \frac{\pi}{\beta T} = Q. \quad (24.115)$$

Иными словами, добротность  $Q$  характеризует как затухание, так и резонансные свойства колебательной системы.

## Контрольные вопросы

1. Перечислите основные характеристики гармонического колебательного движения.
2. Чему равны период и частота малых колебаний для шарика, подвешенного на пружине?
3. Чему равны период и частота малых колебаний для математического маятника?
4. Как будут идти маятниковые часы, если их перенести с Земли на Луну? Тот же вопрос для часов с пружинным маятником.
5. Как скажется на ходе маятниковых часов смена времен года, если они установлены в неотапливаемой комнате?



6. Изменится ли период колебаний качелей, если на них вместо одного человека сядут двое?
7. Как надо передвинуть чечевицу маятниковых часов, если часы отстают?
8. Чему равен период колебаний математического маятника в свободно падающем лифте?
9. Изменится ли период колебаний математического маятника, подвешенного в вагоне поезда, если поезд начнет двигаться с ускорением?
10. Миноискатель представляет собой генератор незатухающих электромагнитных колебаний. Индуктивность выполнена в виде проволочного кольца, приближение которого к металлическому предмету вызывает в наушниках смену тона — с высокого на низкий. Объясните явление.
11. Что происходит с энергией при колебательном движении?
12. Если частица совершает гармонические колебания с амплитудой  $A$ , то какое расстояние она пройдет за один период? Каково будет при этом перемещение частицы?
13. Могут ли в какой-то момент совпасть направления векторов смещения и скорости пружинного маятника? Могут ли совпасть направления векторов смещения и ускорения?
14. Тело массой  $m$  подвешено на пружине жесткостью  $k$ . Пружину разрезают на две равные части. Как изменится частота колебаний, если то же тело подвесить на одной из половинок пружины?
15. Как изменится полная механическая энергия колебания, если удвоить его амплитуду?
16. Два тела с одинаковыми массами подвешены на двух одинаковых пружинах. Одно тело оттягивают вниз на расстояние  $a_1 = 10$  см, другое — на расстояние  $a_2 = 20$  см и одновременно отпускают. Какое из них первым пройдет положение равновесия?
17. Тело массой  $m$  колеблется на пружине жесткостью  $k$  с амплитудой  $A$ . Используя анализ размерностей, найдите период колебаний тела.
18. Складываются два одинаково направленных гармонических колебания. Периоды  $T$  и амплитуды  $A$  колебаний равны, а фазы смещены относительно друг друга на  $\pi/2$ . Какова амплитуда  $A_p$  результирующего колебания? Какова разность фаз между результирующим колебанием и составляющими колебаниями?
19. Сложите графически два одинаково направленных гармонических колебания равных периодов, но смещенные по фазе относительно друг друга на  $\pi$ . Амплитуды составляющих колебаний относятся между собой как 3:1.
20. Что такое биения?
21. Что происходит с частотами двух одинаковых пружинных маятников, если их связать мягкой пружиной?
22. Будут ли гармоническими колебания связанных осцилляторов? Что такое нормальные моды?
23. Как изменится период колебаний маятника, если его переместить из воздуха в вязкую среду — воду или масло?
24. Что такое резонанс? Приведите примеры из повседневной жизни.

## Глава 25

# Переменный ток

### 25.1 Квazистационарные токи

Закон Ома и вытекающие из него правила Кирхгофа справедливы для постоянного тока. Однако закон Ома и, следовательно, правила Кирхгофа в некоторых случаях могут быть применены и для переменного тока, если мгновенные значения изменяющихся во времени токов и напряжений за время  $\tau = l/c$ , необходимое для передачи электромагнитного возмущения в любую точку цепи длиной  $l$ , изменяются столь незначительно, что их можно считать во всех участках цепи практически одинаковыми. Такие переменные токи называются *квazистационарными*. Условие квazистационарности будет выполнено, если  $\tau = l/c \ll T$ , где  $T$  — период изменения величины тока.

Например, для переменного тока промышленной частоты ( $\nu = 50$  Гц) имеем  $cT = c/\nu = 3 \cdot 10^8/50 = 6 \cdot 10^6$  м = 6000 км. Принимая, что длина цепи должна составлять не более нескольких процентов от  $cT$ , получаем, что ток квazистационарен для цепей длиной не более 100 км. Следовательно, в пределах этих размеров мгновенные значения подчиняются закону Ома и правилу Кирхгофа. При изучении разделов, предлагаемых ниже, предполагается, что рассматриваемые в них токи квazистационарны.

### 25.2 Переменный ток через элементы цепи

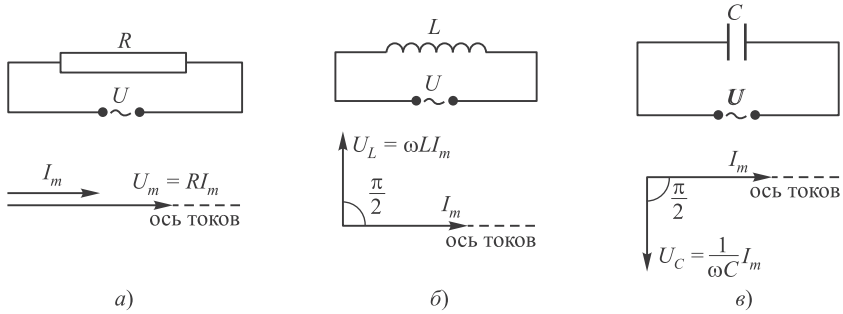
#### Переменный ток через сопротивление

Если к активному сопротивлению приложено переменное напряжение

$$U = U_m \cos \omega t, \quad (25.1)$$

то по закону Ома текущий через это сопротивление ток будет равен:

$$I = \frac{U}{R} = \frac{U_m}{R} \cos \omega t = I_m \cos \omega t. \quad (25.2)$$



**Рис. 25.1.** Переменный ток через активное сопротивление (а), индуктивность (б) и емкость (в)

Следовательно, связь между амплитудами силы тока и напряжения в цепи выражается соотношением

$$I_m = \frac{U_m}{R}. \quad (25.3)$$

Из (25.3) следует, что колебания напряжения и тока на активном сопротивлении происходят в одной фазе, что иллюстрирует *векторная диаграмма* колебаний (рис. 25.1, а), на которой соответствующие векторы параллельны.

### Переменный ток через индуктивность

Пусть переменное напряжение, изменяющееся по гармоническому закону, включено в цепь с индуктивностью  $L$ , не обладающей ни емкостью, ни сопротивлением. В этом случае в цепи должна возникнуть ЭДС самоиндукции  $\mathcal{E}_i = -LdI/dt$ , направленная против ЭДС источника тока. Поскольку активное сопротивление  $R$  в цепи равно нулю, закон Ома запишется в виде

$$U + \mathcal{E}_i = IR = 0, \quad \text{или} \\ L \frac{dI}{dt} = U_m \cos \omega t. \quad (25.4)$$

Решение этого дифференциального уравнения имеет вид

$$I = \frac{U_m}{\omega L} \sin \omega t + \text{const}. \quad (25.5)$$

Поскольку в цепи действует лишь переменное напряжение и нет другого источника тока, его постоянная составляющая равна нулю:

$$I = \frac{U_m}{\omega L} \sin \omega t = \frac{U_m}{\omega L} \cos \left( \omega t - \frac{\pi}{2} \right) = \\ = I_m \cos \left( \omega t - \frac{\pi}{2} \right), \quad (25.6)$$

где  $U_m/\omega L = I_m$ . Сопоставляя полученное выражение с законом Ома для постоянного тока, замечаем, что роль сопротивления играет произведение  $\omega L$ .

Эта величина называется *индуктивным сопротивлением* и обозначается как  $X_L$ :

$$X_L = \omega L. \quad (25.7)$$

Следовательно, величина индуктивного сопротивления растет с частотой  $\omega$ . Постоянному току ( $\omega = 0$ ) индуктивность сопротивления не оказывает.

В данном случае напряжение  $U_L$  на индуктивности совпадает с напряжением, вырабатываемым источником тока. Видно, что напряжение опережает по фазе на  $\pi/2$  ток через индуктивность. Векторная диаграмма изображена на рис. 25.1, б.

### Переменный ток через емкость

Пусть переменное напряжение  $U = U_m \cos \omega t$  включено в цепь с емкостью  $C$ . Предположим, что индуктивностью и сопротивлением подводящих проводов можно пренебречь. Поскольку емкость непрерывно перезаряжается, в цепи будет идти переменный ток. При отсутствии падения напряжения в проводах напряжение на конденсаторе равно внешнему напряжению:

$$U_C = \frac{q}{C} = U_m \cos \omega t. \quad (25.8)$$

Так как  $I = dq/dt$ , получаем, что

$$\begin{aligned} I &= -\omega C U_m \sin \omega t = \omega C U_m \cos \left( \omega t + \frac{\pi}{2} \right) = \\ &= I_m \cos \left( \omega t + \frac{\pi}{2} \right), \end{aligned} \quad (25.9)$$

где

$$I_m = \omega C U_m = \frac{U_m}{1/\omega C} = \frac{U_m}{X_C}. \quad (25.10)$$

Величина

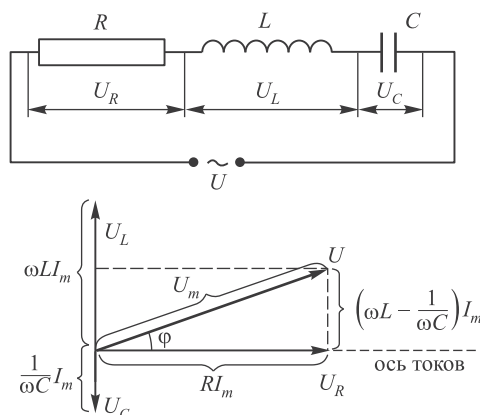
$$X_C = \frac{1}{\omega C} \quad (25.11)$$

называется *емкостным сопротивлением*. Очевидно, что через емкость может идти переменный ток тем больший, чем больше частота его изменения  $\omega$  и емкость конденсатора  $C$ . Для постоянного тока  $\omega = 0$  и емкостное сопротивление становится бесконечно большим: постоянный ток не может идти через конденсатор.

Напряжение на емкости  $U_C$  совпадает с внешним напряжением. Видно, что напряжение на емкости отстает по фазе от тока на  $\pi/2$ , что иллюстрирует векторная диаграмма (рис. 25.1, в).

## 25.3 Цепь переменного тока

Цепь переменного тока, содержащая емкость, индуктивность и активное сопротивление, показана на рис. 25.2. В цепь включен источник переменного напряжения  $U = U_m \cos \omega t$ . В такой цепи будет идти переменный ток, параметры которого определяются величинами  $R$ ,  $L$ ,  $C$ . Внешнее напряжение должно быть равно векторной сумме падений напряжения на отдельных элементах цепи. Эту сумму найдем с помощью векторной диаграммы.



**Рис. 25.2.** Переменный ток через цепь, содержащую активное сопротивление, емкость и индуктивность

Обозначим амплитуду тока через  $I_m$ . На активном сопротивлении  $R$  ток вызовет падение напряжения с амплитудой  $U_R = I_m R$ , причем фаза напряжения совпадает с фазой тока. На векторной диаграмме вектор  $U_R$  нужно отложить вдоль оси  $x$  — оси токов (см. рис. 25.2). Падение напряжения на индуктивности с амплитудой  $U_L = \omega L I_m$  опережает ток по фазе на  $\pi/2$  и на векторной диаграмме будет отложено в положительном направлении оси  $y$ . Падение напряжения на емкости будет иметь амплитуду  $U_C = I_m / (\omega C)$  и фазу, отстающую на  $\pi/2$  от фазы тока. На векторной диаграмме вектор  $U_C$  отложен в отрицательном направлении оси  $y$ .

Падения напряжения  $U_R$ ,  $U_L$ ,  $U_C$  в сумме должны быть равны приложенному к цепи напряжению  $U$ . Поэтому, сложив векторы, изображающие  $U_R$ ,  $U_L$ ,  $U_C$ , получим вектор  $U$  длиной  $U_m$  (см. рис. 25.2). Из рисунка видно, что  $U_m^2 = U_R^2 + (U_L - U_C)^2$ , или

$$U_m^2 = R^2 I_m^2 + \left[ \left( \omega L - \frac{1}{\omega C} \right) I_m \right]^2, \quad (25.12)$$

откуда

$$I_m = \frac{U_m}{\sqrt{R^2 + \left( \omega L - \frac{1}{\omega C} \right)^2}}. \quad (25.13)$$

При этом разность фаз между током в цепи и напряжением составляет угол  $\varphi$ , который определится условием (см. рис. 25.2)

$$\operatorname{tg} \varphi = \frac{\left( \omega L - \frac{1}{\omega C} \right)}{R}. \quad (25.14)$$

Поскольку ток отстает по фазе от напряжения, закон изменения тока имеет вид

$$I = I_m \cos(\omega t - \varphi). \quad (25.15)$$

В сущности, мы другим методом воспроизвели решения для вынужденных колебаний, полученные выше. В самом деле, подставляя значения  $\omega_0 = 1/\sqrt{LC}$ ,  $\beta = R/2L$ ,  $a_m = U_m/L$  в выражение (24.109) для амплитуды колебаний заряда на емкости, получаем

$$A_Q = \frac{U_m}{\omega \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}. \quad (25.16)$$

Подставляя те же значения в выражение (24.110) для фазы колебаний заряда на емкости, находим:

$$\operatorname{tg} \varphi_Q = \frac{2\beta\omega}{\omega^2 - \omega_0^2} = \frac{R}{\omega L - \frac{1}{\omega C}}. \quad (25.17)$$

Решение (24.108) для колебаний заряда на емкости имеет вид

$$Q = A_Q \cos(\omega t + \varphi_Q).$$

Дифференцируя по времени, находим ток в цепи:

$$I = \frac{dQ}{dt} = -\omega A_Q \sin(\omega t + \varphi_Q) = \omega A_Q \cos\left(\omega t + \frac{\pi}{2} + \varphi_Q\right). \quad (25.18)$$

Сравнивая с выражением (25.15), убеждаемся, что амплитуда тока определяется выражением  $I_m = \omega A_Q$ , которое совпадает с (25.13). Для начальной фазы колебаний тока  $\varphi = -\pi/2 - \varphi_Q$ , откуда

$$\operatorname{tg} \varphi = \operatorname{ctg} \varphi_Q = \frac{1}{\operatorname{tg} \varphi_Q},$$

что совпадает с (25.14). Используя решения для вынужденных колебаний, мы обосновали правомерность применения метода векторных диаграмм в цепи переменного тока.

Итак, если напряжение, поданное на цепь, содержащую последовательно включенные  $R$ ,  $L$  и  $C$ , изменяется по закону  $U = U_m \cos \omega t$ , то в такой цепи течет переменный ток  $I = I_m \cos(\omega t - \varphi)$ , где величины  $I_m$  и  $\varphi$  определяются формулами (25.13) и (25.14) соответственно. Величина

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2} \quad (25.19)$$

называется **полным сопротивлением цепи**, а величина  $\left(\omega L - \frac{1}{\omega C}\right)$  называется **реактивным сопротивлением**. При  $\omega L > 1/(\omega C)$  ток отстает по фазе от напряжения ( $\varphi > 0$ ), при  $\omega L < 1/(\omega C)$  опережает его ( $\varphi < 0$ ).

## 25.4 Резонансные явления

Изучая вынужденные колебания, мы видели, что при  $\omega = \omega_{\text{рез}}$  амплитуда колебаний заряда резко увеличивается. Резонансная частота меньше собственной

частоты колебаний в контуре:

$$\omega_{\text{рез}} = \sqrt{\omega_0^2 - 2\beta^2} = \sqrt{\frac{1}{LC} - \frac{R^2}{2L^2}} < \frac{1}{\sqrt{LC}} = \omega_0. \quad (25.20)$$

При этой же частоте максимума достигает амплитуда колебаний напряжения на конденсаторе  $U_C = Q/C$ . Выражение для тока получается дифференцированием решения  $Q(t)$  для заряда, что дает в амплитуде дополнительную зависимость от частоты  $\omega$ . Поэтому резонансная частота для тока отличается, как мы увидим ниже, от  $\omega_{\text{рез}}$ .

## Резонанс напряжений

Как следует из (25.13), амплитуда тока достигает максимума при таком значении внешней частоты, когда обращается в нуль реактивное сопротивление:  $\omega L - \frac{1}{\omega C} = 0$ , откуда  $\omega = 1/\sqrt{LC} = \omega_0$ . Тогда угол  $\varphi$  равен 0, а изменение тока и приложенного напряжения происходят синхронно. В этом случае полное сопротивление  $Z = Z_{\min} = R$  и амплитуда тока достигает максимального значения  $I_{m,\text{рез}} = U_m/R$ . При этом падения напряжения на емкости и индуктивности равны по величине ( $U_C = U_L$ ) и противоположны по фазе, так что они взаимно компенсируются. Это явление называется **резонансом напряжений**.

Амплитуда напряжения на конденсаторе при таком резонансе равна:

$$U_{C,\text{рез}} = \frac{I_{m,\text{рез}}}{\omega_0 C} = \frac{1}{R} \sqrt{\frac{L}{C}} U_m. \quad (25.21)$$

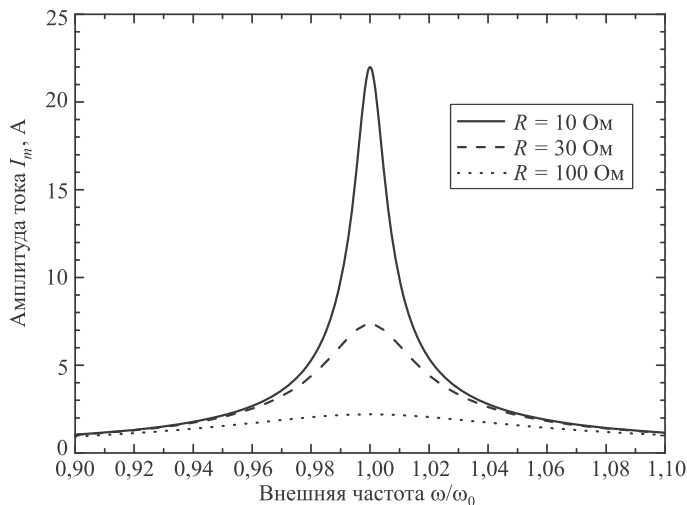
Мы получили, что при резонансе напряжений отношение амплитуды напряжения на конденсаторе к амплитуде внешнего напряжения равно добротности контура. Резонансная кривая для тока (рис. 25.3) отличается от резонансной кривой для заряда (см. рис. 24.20).

## Резонанс токов

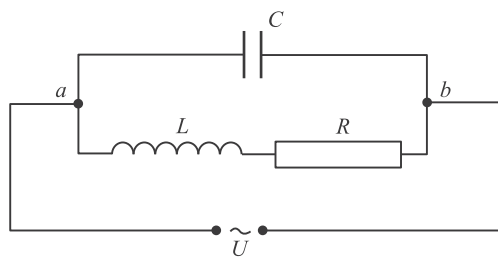
В предыдущем разделе мы включили генератор внешнего напряжения последовательно в электромагнитный контур. Сейчас мы займемся резонансом в случае, когда контур подключен к генератору параллельно. Рассмотрим цепь, состоящую из двух параллельных ветвей, одна из которых содержит емкость  $C$ , а другая — индуктивность  $L$  и активное сопротивление  $R$  (рис. 25.4).

В случае простой неразветвленной цепи (см. рис. 25.2) для всех элементов цепи ( $L$ ,  $C$ ,  $R$ ) величина силы тока была одной и той же и задача сводилась к сложению величин напряжения на индуктивности, емкости и сопротивлении. В рассматриваемом случае (см. рис. 25.4) общим является напряжение между точками  $a$  и  $b$  для обеих ветвей цепи, а силы тока в ветвях  $I_C$  и  $I_L$  различны. Согласно правилу Кирхгофа, полная сила тока  $I = I_C + I_L$ , поэтому задача сводится к сложению токов (рис. 25.5).

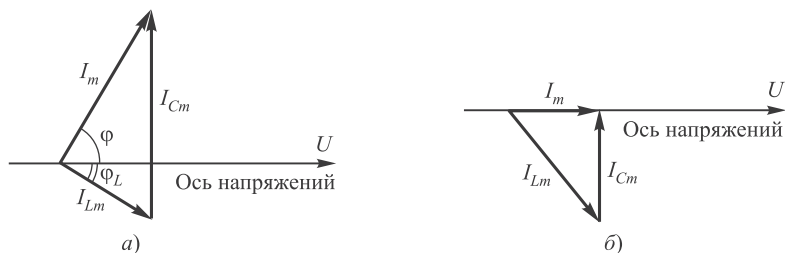
Если мы изобразим колебания напряжения между точками  $a$  и  $b$  векторами, направленными вдоль линии  $U$  (ось напряжений), то колебания тока в цепи,



**Рис. 25.3.** Резонансные кривые для амплитуды переменного тока при резонансе напряжений построены для значений  $U_m = 220$  В,  $C = 100$  пФ,  $L = 100$  мкГн и различных значений активного сопротивления  $R$ . При таких параметрах цепи резонансное значение частоты  $\omega_0 = 1/\sqrt{LC} = 10$  МГц. При малых ( $\omega \ll \omega_0$ ) и больших ( $\omega \gg \omega_0$ ) частотах ток стремится к нулю. Положение максимума  $\omega = \omega_0$  не зависит от затухания в цепи



**Рис. 25.4.** Электрическая цепь для наблюдения резонанса токов



**Рис. 25.5.** Векторная диаграмма для токов в ветвях параллельно подключенного электромагнитного контура (а); при резонансе токов (б) контур ведет себя как активное сопротивление: сдвиг фаз между результирующим током и напряжением отсутствует



содержащей индуктивность и активное сопротивление, изобразятся вектором  $I_{Lm}$  (рис. 25.5, а), длина которого равна:

$$I_{Lm} = \frac{U_m}{\sqrt{R^2 + \omega^2 L^2}}, \quad (25.22)$$

а угол его наклона  $\varphi_L$  к оси напряжений можно найти из условия

$$\operatorname{tg} \varphi_L = \frac{\omega L}{R}. \quad (25.23)$$

При отсутствии конденсатора полный ток в цепи с индуктивностью равен  $I_{Lm}$  и отстает по фазе от напряжения, что видно из рис. 25.5.

Колебания тока в другой ветви цепи, содержащей емкость  $C$ , изобразятся вектором  $I_{Cm}$ . При чисто емкостном сопротивлении ток опережает по фазе внешнее напряжение на  $\pi/2$ . Соответствующий вектор на рис. 25.5 повернут относительно оси напряжений на угол  $+\pi/2$ , а его длина (амплитуда) равна:

$$I_{Cm} = U_m \omega C. \quad (25.24)$$

Колебания полного тока  $I$  определяются векторной суммой токов  $I_{Cm}$  и  $I_{Lm}$ , т. е. вектором  $I_m$ .

Если в цепи, изображенной на рис. 25.4, изменить  $L$ ,  $C$ ,  $R$  и  $\omega$ , то при некотором соотношении между этими величинами сдвиг фаз  $\varphi$  между полным током и напряжением станет равным нулю и, следовательно, контур будет вести себя как чисто активное сопротивление (рис. 25.5, б). Этот частный случай называется **резонансом токов**. При резонансе полный ток становится минимальным. Следовательно, при резонансе токов сопротивление контура достигает наибольшего значения. В отличие от случая резонанса напряжений оно не равно активному сопротивлению в цепи  $R$ , а зависит от величин  $L$  и  $C$ .

Найдем условие, при котором наступает резонанс токов. Из рисунка 25.5, б следует, что

$$I_{Cm} = I_{Lm} \sin \varphi_L, \quad (25.25)$$

где

$$\sin \varphi_L = \frac{\omega L}{\sqrt{R^2 + \omega^2 L^2}}. \quad (25.26)$$

Следовательно, можно записать:

$$U_m \omega C = \frac{\omega L}{\sqrt{R^2 + \omega^2 L^2}} \cdot \frac{U_m}{\sqrt{R^2 + \omega^2 L^2}} = U_m \omega \frac{L}{R^2 + \omega^2 L^2},$$

откуда

$$\omega_{\text{рез}} = \sqrt{\frac{1}{LC} - \frac{R^2}{L^2}}. \quad (25.27)$$

При малом затухании колебаний ( $R \ll \sqrt{L/C}$ ) частота  $\omega_{\text{рез}} \simeq \sqrt{\frac{1}{LC}} = \omega_0$ , т. е. в этом случае для резонанса токов (как и для резонанса напряжений) частота колебаний внешнего напряжения должна совпадать с частотой  $\omega_0$  собственных колебаний контура.

## 25.5 Мощность в цепи переменного тока

Мгновенное значение мощности, выделяемой в цепи переменного тока, равно:

$$\begin{aligned} P(t) &= U(t)I(t) = U_m \cos \omega t \cdot I_m \cos(\omega t - \varphi) = \\ &= \frac{1}{2} U_m I_m (\cos \varphi + \cos(2\omega t - \varphi)). \end{aligned} \quad (25.28)$$

При усреднении за период колебаний второе слагаемое в скобках пропадает, и мы получаем

$$P = \langle P(t) \rangle = \frac{U_m I_m}{2} \cos \varphi. \quad (25.29)$$

Используя (25.14), можем записать:  $\cos \varphi = R/Z$ , где  $Z$  — полное сопротивление цепи (25.19). Но амплитуды тока и напряжения связаны соотношением (25.13):  $U_m = I_m Z$ . Тогда (25.29) можно представить в другом виде:

$$P = \frac{I_m^2 R}{2}. \quad (25.30)$$

Такую же мощность развивает постоянный ток

$$I_{\text{эф}} = \frac{I_m}{\sqrt{2}}. \quad (25.31)$$

Величину  $I_{\text{эф}}$  называют *эффективной* или *среднеквадратичной силой тока*. Аналогично определяется *эффективное (среднеквадратичное) напряжение*:

$$U_{\text{эф}} = \frac{U_m}{\sqrt{2}}, \quad (25.32)$$

так что среднюю мощность (25.29) можно записать как

$$P = U_{\text{эф}} I_{\text{эф}} \cos \varphi. \quad (25.33)$$

Когда мы говорим в быту о напряжении 220 В, то это относится именно к среднеквадратичному напряжению. Максимальное значение напряжения в цепи равно  $U_m = \sqrt{2} \times 220 \text{ В} = 311 \text{ В}$ .

Величину  $\cos \varphi$  в (25.33) называют *коэффициентом мощности*. Чтобы добиться выделения в цепи возможно большей мощности, надо сделать сдвиг фаз между током и напряжением в цепи близким к нулю. Например, если слишком велика индуктивность цепи, в нее добавляют конденсаторы.

## Контрольные вопросы

1. Что такое квазистационарный ток? При каких условиях токи можно считать квазистационарными?
2. Чем различаются переменные токи, текущие через активное сопротивление, емкость и индуктивность?
3. Что такое реактивное сопротивление? Чему равно полное сопротивление?
4. Какие резонансные явления наблюдаются в цепи переменного тока?
5. Постройте векторные диаграммы и рассмотрите резонанс токов для контура, где активное сопротивление и конденсатор включены последовательно, а индуктивность — параллельно им (т.е. для контура, подобного изображенному на рис. 25.4, с той разницей, что контакт  $b$  перенесен в точку между индуктивностью  $L$  и активным сопротивлением  $R$ ).
6. Что такое средняя мощность в цепи переменного тока, среднеквадратичные значения тока и напряжения?

## Глава 26

# Нелинейные колебания

Колебания осуществляются не только в простых системах типа маятников и электрических колебательных контуров. Как правило, ни одна система в природе не находится в стационарном состоянии, но испытывает колебания вокруг тех или иных равновесных значений. Для простоты будем рассуждать об одномерной системе. Если ее потенциальная энергия  $U(x)$  имеет минимум в точке  $x = x_0$ , то эта точка соответствует положению равновесия системы. При малых отклонениях от положения равновесия ( $|x - x_0| \ll x_0$ ) потенциальную энергию можно разложить в ряд Тейлора, ограничиваясь его первыми членами:

$$U(x) \approx U(x_0) + (x - x_0)U'(x_0) + \frac{(x - x_0)^2}{2}U''(x_0).$$

В точке устойчивого равновесия первая производная потенциальной энергии равна нулю, а вторая производная положительна:  $k = U''(x_0) > 0$ . Поэтому для малых отклонений от положения равновесия потенциальная энергия имеет тот же вид, что и в линейных системах:

$$U(x) \approx U(x_0) + \frac{kx^2}{2}.$$

Сила, действующая на частицы в потенциальном поле, равна взятому с обратным знаком градиенту потенциальной энергии; в рассматриваемом нами одномерном случае

$$F = -\frac{dU(x)}{dx} \approx -kx.$$

Соответственно уравнение второго закона Ньютона  $m\ddot{x} = F$  совпадает с уравнением колебаний гармонического маятника:

$$\ddot{x}(t) + \omega_0^2 x(t) = 0,$$

где  $\omega_0 = \sqrt{k/m}$ .

Но мы пришли к этому уравнению в предположении малости колебаний, что в природе осуществляется далеко не всегда. При больших отклонениях от положения равновесия становятся важны следующие члены разложения потенциальной энергии. Практически в любой системе имеется так называемый *нелинейный элемент*, наличие которого приводит к тому, что колебания с большой амплитудой

перестают быть гармоническими. За примерами далеко ходить не надо. Если мы говорим о пружинном маятнике, то следует учесть приближенный характер закона Гука  $F = -kx$ . При больших растяжениях пружины возникают остаточные деформации, а в какой-то момент может произойти и ее разрушение. При колебаниях маятника в вязкой среде сила сопротивления зависит от скорости линейно лишь при небольших скоростях движения (ламинарное течение), а при увеличении скорости в игру вступает сила лобового сопротивления, пропорциональная квадрату скорости. В электромагнитном колебательном контуре катушки индуктивности наматываются на сердечник из ферромагнетика, в котором при сильных полях возникает нелинейное явление гистерезиса. Таким образом, изученные нами гармонические колебания — предельный (и самый простой) случай колебаний, встречающихся в природе.

## 26.1 Нелинейные колебания маятника

Мы познакомимся с основными свойствами нелинейных колебаний на примере маятника, представляющего собой груз массой  $m$ , подвешенный на невесомой и нерастяжимой штанге длиной  $l$ , которая может свободно и без трения вращаться вокруг точки подвеса. Движение такого маятника описывается уравнением (24.5)

$$\ddot{\varphi} = -\omega_0^2 \sin \varphi, \quad \omega_0 = \sqrt{\frac{g}{l}},$$

где  $\varphi$  — угол отклонения штанги маятника от вертикали. Отклоним наш маятник на угол  $\varphi_0$  и отпустим без начальной скорости. Иными словами, мы будем решать уравнение движения маятника при начальных условиях  $\varphi(0) = \varphi_0, \dot{\varphi}(0) = 0$ . Метод решения стандартный: умножаем уравнение движения на скорость  $\dot{\varphi}(t)$  и замечаем, что слева и справа от знака равенства стоят полные производные:

$$\frac{d}{dt} \frac{\dot{\varphi}^2}{2} = \omega_0^2 \frac{d}{dt} \cos \varphi. \quad (26.1)$$

Интегрируя и используя начальные условия, получаем закон сохранения энергии

$$\dot{\varphi}^2 = 2\omega_0^2 (\cos \varphi - \cos \varphi_0). \quad (26.2)$$

Для определения периода колебаний  $T$  маятника достаточно рассмотреть его движение от начального положения  $\varphi = \varphi_0$  до момента прохождения точки равновесия  $\varphi = 0$ . На этом участке пути уравнение (26.2) можно записать в виде

$$\dot{\varphi} = -\omega_0 \sqrt{2(\cos \varphi - \cos \varphi_0)}. \quad (26.3)$$

Интегрируя, получаем уравнение, определяющее зависимость угла отклонения  $\varphi$  маятника от времени:

$$\frac{t}{T_0} = \frac{1}{2\pi} \int_{\varphi}^{\varphi_0} \frac{dx}{\sqrt{2(\cos x - \cos \varphi_0)}}. \quad (26.4)$$

Здесь  $T_0 = 2\pi/\omega_0$  — период гармонических колебаний нашего маятника при малых амплитудах.

За четверть периода  $t = T/4$  маятник опустится в вертикальное положение ( $\varphi = 0$ ). Тогда из (26.4) получаем зависимость периода колебаний  $T$  от их амплитуды  $\varphi_0$ :

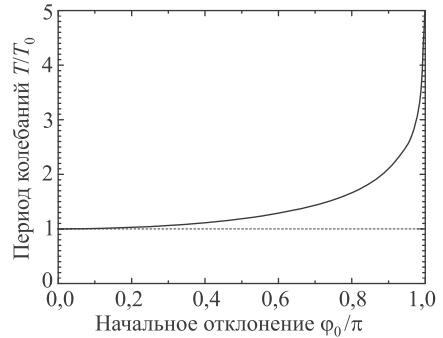
$$\frac{T}{T_0} = \frac{\sqrt{2}}{\pi} \int_0^{\varphi_0} \frac{dx}{\sqrt{\cos x - \cos \varphi_0}}. \quad (26.5)$$

Теперь легко определить значения периода колебаний при разных углах начального отклонения маятника. При малых углах период  $T$  близок к периоду  $T_0$  гармонических колебаний, но при увеличении угла  $\varphi_0$  период  $T$  растет. Нетрудно сообразить, почему это происходит: возвращающая сила пропорциональна синусу угла  $\varphi$ , а  $\sin \varphi < \varphi$ , т.е. эта сила меньше той, что действует на маятник в линейном приближении. Соответственно маятник медленнее набирает скорость при движении к положению равновесия и ему потребуется больше времени, чтобы его достичь. Так, из (26.5) следует, что при начальном угле отклонения  $\varphi_0 = 3\pi/4$  период нелинейных колебаний в полтора раза больше периода гармонических колебаний того же маятника ( $T = 1,528T_0$ ).

Результаты численного интегрирования уравнения (26.4) представлены на рис. 26.1. Сплошная линия описывает нелинейные колебания маятника в зависимости от начального угла отклонения  $\varphi_0 = 3\pi/4$ . При приближении начального угла отклонения к  $\pi$  период стремится к бесконечности. Это периодическое движение не описывается одной тригонометрической функцией и потому не является гармоническим колебанием. На том же графике штриховая линия показывает гармонические колебания этого же маятника при малых отклонениях, когда период колебаний равен  $T_0$  — мы видим существенные отличия от колебаний с большой амплитудой. На данном конкретном примере мы встретились с двумя фундаментальными свойствами нелинейных колебаний:

- период нелинейных колебаний обычно зависит от их амплитуды (такие колебания называются **неизохронными**, в отличие от изохронных гармонических колебаний);
- нелинейные колебания не являются, вообще говоря, гармоническими (в этом случае говорят, что они **ангармонические**).

Примером гармонических, но неизохронных колебаний может служить движение планет по эллипсам, в фокусе которых находится Солнце. Оно описывается тригонометрической функцией, так что проекция планеты на любую ось, лежащую в плоскости ее орбиты, совершает гармонические колебания. Но, согласно третьему закону Кеплера (см. т. 1, гл. 6), квадраты периодов  $T$  обращения планет вокруг Солнца относятся как кубы больших полуосей  $A$  их орбит, т.е. период обращения планеты зависит от амплитуды ее движения по закону  $T \propto A^{3/2}$ .



**Рис. 26.1.** Зависимость периода нелинейных колебаний маятника от начального угла отклонения

Другим примером является нелинейный осциллятор с потенциальной энергией, составленной из двух половинок парабол гармонических осцилляторов с разными частотами:

$$U(x) = \begin{cases} \frac{m\omega_1^2 x^2}{2}, & x \geq 0; \\ \frac{m\omega_2^2 x^2}{2}, & x \leq 0. \end{cases} \quad (26.6)$$

Очевидно, что период колебаний такого осциллятора равен сумме двух полупериодов каждого из гармонических осцилляторов и не зависит от амплитуды колебаний:

$$T = \frac{\pi}{\omega_1} + \frac{\pi}{\omega_2}.$$

Эти колебания явно ангармонические, но в то же время изохронные.

Говоря об изохронности малых колебаний, мы предполагали, что в точке равновесия отлична (и положительна) вторая производная потенциальной энергии. Если это не так, даже малые колебания такой системы не будут изохронными. Примером может служить так называемый **ангармонический осциллятор**, потенциальная энергия которого представляется в виде  $U(x) = k|x|^n/n$ ,  $n > 0$ . Пусть в такой потенциальной яме колеблется частица массой  $m$ , причем начальные условия имеют вид  $x(0) = x_0$ ,  $\dot{x}(0) = 0$ . Закон сохранения энергии (или, как говорят, первый интеграл уравнения движения) имеет вид

$$\frac{m\dot{x}^2}{2} + \frac{k|x|^n}{n} = \frac{k|x_0|^n}{n}. \quad (26.7)$$

Разделяя переменные и интегрируя по первой четверти периода колебаний, находим связь периода  $T$  с амплитудой  $x_0$ :

$$\sqrt{\frac{m}{2}} \int_0^{x_0} \frac{dx}{\sqrt{x_0^n - x^n}} = \frac{T}{4} \sqrt{\frac{k}{n}}. \quad (26.8)$$

Заменяя переменную в интеграле, приходим к выражению для периода колебаний ангармонического осциллятора:

$$T = 4 \sqrt{\frac{nm}{2kx_0^{n-2}}} \int_0^1 \frac{dx}{\sqrt{1-x^n}}. \quad (26.9)$$

При  $n = 2$  интеграл в (26.9) равен  $\pi/2$ , и мы получаем известную формулу для периода колебаний гармонического осциллятора на пружине жесткостью  $k$ :  $T = 2\pi\sqrt{m/k}$ . При  $n \neq 2$  период степенным образом зависит от амплитуды, так что колебания (неважно, малые или большие) не являются изохронными.

Любое ангармоническое колебание  $x(t)$  с периодом  $T$  можно разложить в ряд Фурье, т. е. представить как сумму гармонических колебаний с частотами, кратными основной частоте:

$$x(t) = A_0 + \sum_{i=1}^{\infty} A_i \cos(i\omega t + \varphi_i). \quad (26.10)$$

Нулевой член всегда можно убрать, сместив начало отсчета переменной  $x(t)$ . Колебание с частотой  $i\omega$  называется  $i$ -й гармоникой. Обычно в сигнале доминирует первая гармоника с самой большой амплитудой  $|A_1|$ . На рисунке 26.2 кривая из точек соответствует первой гармонике — гармоническому колебанию с периодом  $T = 1,528T_0$ . Видно, что колебания нашего маятника (сплошная кривая) не являются гармоническими, но их отличия от первой гармоники невелики.

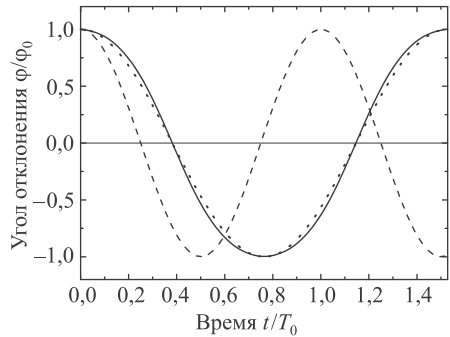
Возникновение высших гармоник можно представить себе как результат преобразования гармонического сигнала в нелинейном элементе. Пусть, к примеру, на такой элемент подается сигнал  $x(t)$ , а выходит сигнал  $y(t) = a_1x(t) + a_2x^2(t)$ . Если входной сигнал имеет стандартный вид  $x(t) = x_0 \cos \omega t$ , то выходной сигнал запишется как

$$\begin{aligned} y(t) &= a_1x_0 \cos \omega t + a_2x_0^2 \cos^2 \omega t = a_1x_0 \cos \omega t + a_2x_0^2 \frac{1 + \cos 2\omega t}{2} = \\ &= A_0 + A_1 \cos \omega t + A_2 \cos 2\omega t, \end{aligned} \quad (26.11)$$

где  $A_0 = a_2x_0^2/2$ ,  $A_1 = a_1x_0$ ,  $A_2 = a_2x_0^2/2$ . Мы видим, что нулевой член и вторая гармоника появились в выходном сигнале вследствие квадратичной нелинейности приемного элемента. При кубической нелинейности мы получили бы третью гармонику и добавку к первой и т. д. Практическую реализацию этих простых тригонометрических преобразований можно наблюдать в красивом явлении нелинейной оптики — генерации второй гармоники, когда инфракрасный (невидимый глазу) луч лазера падает на нелинейный кристалл и при достаточно высокой интенсивности излучения оттуда выходит зеленый луч видимого диапазона с удвоенной частотой.

## 26.2 Фазовый портрет

Состояние механической системы можно изобразить точкой в многомерном **фазовом пространстве**, на координатных осях которого отложены значения пространственных координат частиц системы и значения их импульсов (скоростей). Если в систему входит  $N$  частиц, то вследствие трехмерности нашего пространства общее число их пространственных координат равно  $3N$  и столько же координат характеризуют их импульс. Таким образом, фазовое пространство системы  $N$  частиц имеет размерность  $6N$ . В простейшем случае одной частицы, движущейся в одномерном пространстве, ее динамическое состояние характеризуется одной координатой  $x(t)$  и одной скоростью  $\dot{x}(t)$ . Тогда размерность фазового пространства равна двум, и мы говорим о **фазовой плоскости**. Следует отметить также, что слова «координата» и «импульс» надо понимать в обобщенном смысле. Так,



**Рис. 26.2.** Зависимость от времени угла отклонения маятника (сплошная кривая) в сравнении с гармоническими колебаниями того же маятника при малых углах отклонения (штриховая линия) и первой гармоникой (пунктир)

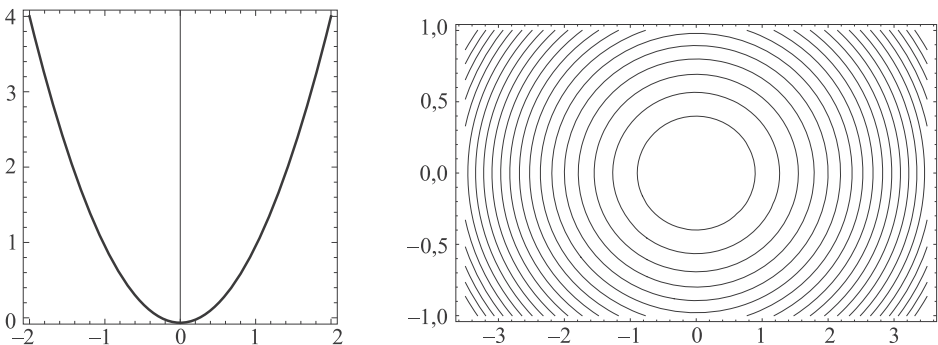


при описании колебаний в электрическом контуре роль координаты играет заряд конденсатора  $q$ , а роль скорости  $\dot{q}$  — идущий в цепи ток.

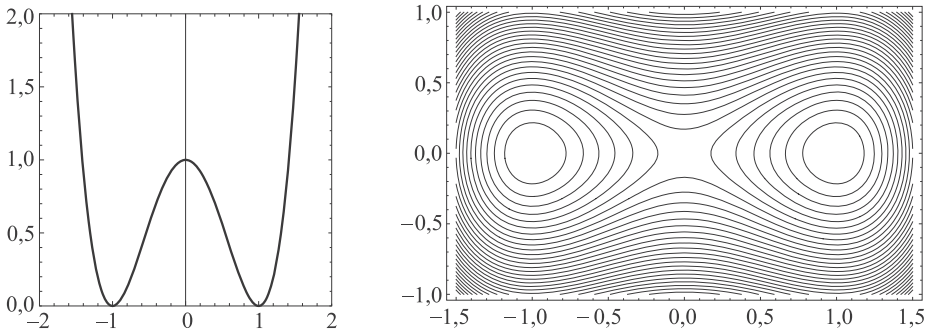
С течением времени точка, отображающая динамическое состояние системы, движется в фазовой плоскости. Для периодического процесса ее траектория является замкнутой линией, так что точка возвращается через каждый период в первоначальное положение, которое задается начальными условиями задачи. Семейство таких кривых, соответствующее разным начальным условиям, называется **фазовым портретом** системы. Так, для гармонического осциллятора траектория точки на фазовой плоскости определяется уравнением закона сохранения энергии  $m\dot{x}^2/2 + kx^2/2 = E$ . При соответствующем выборе единиц измерения расстояний и времени это уравнение превращается в уравнение окружности  $\dot{x}^2 + x^2 = E$ , так что фазовый портрет гармонического осциллятора (рис. 26.3) представляет собой семейство концентрических окружностей с центром в точке  $\{0, 0\}$ , соответствующей состоянию покоя с нулевой энергией. На каждой из таких окружностей полная энергия осциллятора остается постоянной, т. е. траектории в фазовом пространстве являются аналогом изогипс в топографии (см. т. 1, разд. 4.4), изображающих линии постоянной высоты над потенциальной «впадиной».

Так называемый двухямный осциллятор, описываемый потенциальной энергией  $U = (x^2 - 1)^2$ , имеет две «впадины» при  $x = \pm 1$  и седловую точку («перевал») между ними при  $x = 0$ . Сам потенциал и его фазовый портрет изображены на рис. 26.4. Мы видим на нем, что имеются фазовые траектории трех типов. Одни охватывают точку с координатами  $\{-1, 0\}$  и соответствуют колебательному движению осциллятора внутри левой ямы, т. е. около точки  $x = -1$ . Другие охватывают точку с координатами  $\{1, 0\}$  и соответствуют аналогичным колебаниям только в правой яме. Наконец, третье семейство фазовых траекторий охватывает обе ямы. Этот тот случай, когда энергия колебаний превышает высоту потенциального «горба» и в своем движении осциллятор проходит через обе точки равновесия  $x = -1$  и  $x = 1$ .

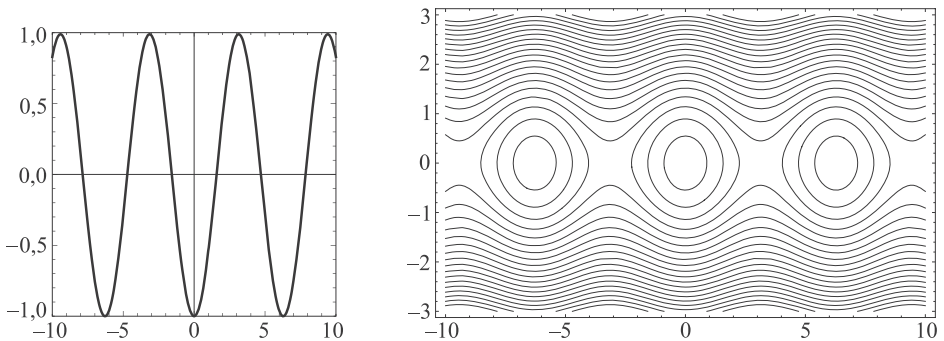
Рассмотрим теперь фазовый портрет нелинейного маятника, потенциальная энергия которого для одного из начальных условий может быть представлена как периодическая функция вида  $U = -\cos x$  с минимумами в точках  $x_n = 2\pi n$ , где



**Рис. 26.3.** Потенциальная энергия гармонического осциллятора (слева) и его фазовый портрет (справа)



**Рис. 26.4.** Потенциальная энергия двухъямного ангармонического осциллятора (слева) и его фазовый портрет (справа)

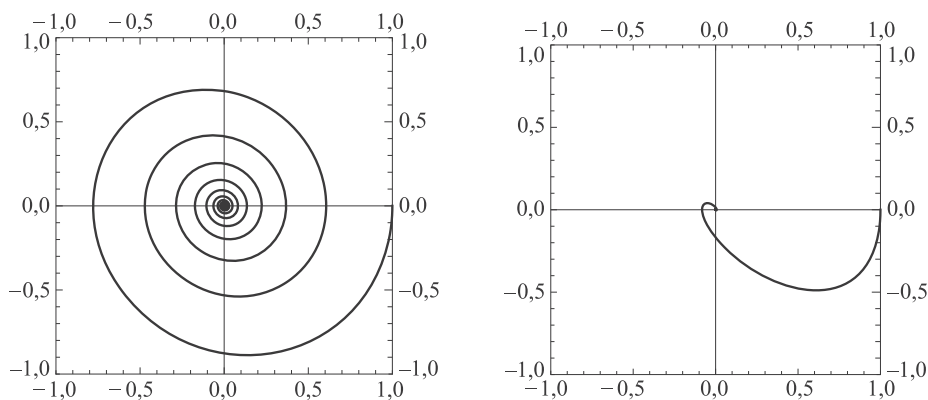


**Рис. 26.5.** Периодическая потенциальная энергия нелинейного осциллятора с минимумами в точках  $x = 0, \pm 2\pi, \pm 4\pi, \dots$  (слева) и его фазовый портрет (справа)

$n = 0, \pm 1, \pm 2, \dots$  Здесь помимо траекторий, охватывающих эти точки минимумов потенциальной энергии, имеются траектории, соответствующие ротационному движению маятника, когда его энергия достаточно велика и он вращается в вертикальной плоскости вокруг точки подвеса, последовательно проходя все точки минимума (рис. 26.5).

Во всех этих случаях вследствие сохранения полной энергии движение осцилляторов было периодическим, а фазовые траектории — замкнутыми линиями. Рассмотрим теперь фазовую траекторию квазигармонического осциллятора с затуханием (см. (24.75)). На рисунке 26.6 представлены фазовые траектории затухающих колебаний при одинаковых начальных условиях, но разной величине коэффициента затухания. В первом случае слабого затухания осциллятор успевает совершить множество колебаний, прежде чем остановится ( $\dot{x} = 0$ ) в точке минимума потенциальной энергии ( $x = 0$ ). Во втором случае начальной энергии осциллятора недостаточно даже для завершения одного полного периода колебаний.

Выбрав иные начальные условия, мы получим спирали, начинающиеся в других точках. Их совокупность и будет фазовым портретом данной системы. Но все эти спирали будут завершаться в центре фазового портрета. Эта точка (с координатами  $\{0, 0\}$  в данном случае), куда при  $t \rightarrow \infty$  стягиваются все фазовые



**Рис. 26.6.** Фазовые траектории осциллятора со слабым (слева) и сильным (справа) затуханием при одинаковых начальных условиях

траектории, называется **аттрактором** (от англ. *attract* — притягивать). Для диссипативной системы характерно «забывание» начальных условий. В общем случае аттрактор — множество точек, посещаемых системой в установившемся режиме, т. е. по прошествии достаточно длительного времени. Для свободных затухающих колебаний аттрактор — одна точка, соответствующая состоянию устойчивого равновесия системы. В следующем разделе мы познакомимся с иной ситуацией, когда аттрактором является некоторая замкнутая кривая — так называемый **предельный цикл**.

### 26.3 Автоколебания

Термины «автоколебания» и «автоколебательные системы» ввел в науку А. А. Андронов (1928). **Автоколебаниями** называют устойчивые незатухающие колебания, характеристики которых — период, амплитуда и форма — определяются параметрами самой системы и **не зависят от начальных условий**.

Автоколебания широко распространены в природе и технике. Когда вы видите полощущийся на ветру флаг, то вы наблюдаете автоколебания. Сходное явление — флаттер крыла самолета или лопасти несущего винта вертолета, возникающий при определенной критической скорости, очень опасно и может привести к разрушению летательного аппарата. Звук, извлекаемый из струнных и духовых музыкальных инструментов, скрип двери, звук человеческого голоса — это тоже автоколебания. Всем известное «рычание» водопроводного крана при определенном темпе расхода воды также служит примером автоколебаний. Это же явление, возникающее в газовых и жидкостных магистралях высокого давления, также может привести к авариям с самыми серьезными последствиями. Автоколебательными системами являются паровые машины, двигатели внутреннего сгорания и электрогенераторы — все устройства, где, в отличие от вынужденных колебаний, происходят незатухающие периодические процессы без воздействия

внешней периодической силы. Обнаружен целый класс автоколебательных химических реакций (так называемые реакции Белоусова—Жаботинского, открытые Б. П. Белоусовым в 1951 г.), в которых, например, определенные растворы могут периодически менять цвет. В автоколебательном режиме функционируют и переменные звезды-цефеиды, периодически меняющие свои размеры и яркость. Благодаря строгой зависимости периода этих пульсаций и светимости звезды они служат в качестве «стандартных свечей» при определении расстояний до галактик (если две цефеиды имеют одинаковые периоды мерцания, но разную видимую светимость, они просто находятся на разном расстоянии от Земли). Наконец, дыхание и работа сердца — это тоже автоколебания.

Независимость характеристик автоколебаний от начальных условий — их главное свойство. Отсюда сразу следует ряд выводов.

- Во-первых, автоколебательная система должна быть диссипативной, т. е. в ней обязательно должно присутствовать затухание, чтобы система не сохраняла память о начальных условиях.
- Во-вторых, колебания автоколебания не затухают, система не может быть замкнутой, изолированной от внешнего мира. Она должна быть открытой и должна быть связана с каким-то источником энергии, который компенсировал бы потери энергии на преодоление сил сопротивления, ведущих к затуханию.
- В-третьих, автоколебательная система должна сама регулировать скорость поступления энергии из внешнего источника, иначе автоколебания не будут устойчивыми — энергии в какой-то момент может не хватить, и колебания затухнут или, наоборот, избыток энергии приведет к безудержному росту колебаний. Это значит, что колебательная система должна иметь обратную связь с источником энергии.
- В-четвертых, в автоколебательной системе обязательно присутствие нелинейного элемента. Амплитуда автоколебаний зависит лишь от самой системы, а в линейной системе она в принципе не может быть определена по очень простой причине. Решение линейных уравнений можно умножить на произвольный множитель, и оно по-прежнему останется решением тех же уравнений. Это означает, что амплитуда не может зависеть от свойств линейной системы, она определяется исключительно начальными условиями.

Исходя из перечисленных свойств автоколебательной системы, мы можем представить себе ее общую блок-схему (рис. 26.7). В нее входят диссипативная колебательная система, источник питания и клапан. Энергия от источника поступает в колебательную систему через клапан — устройство, регулирующее поступление энергии к колебательной системе. В свою очередь, движение самой системы управляет работой клапана.

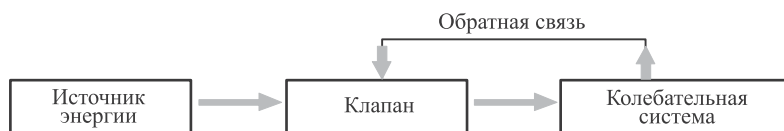
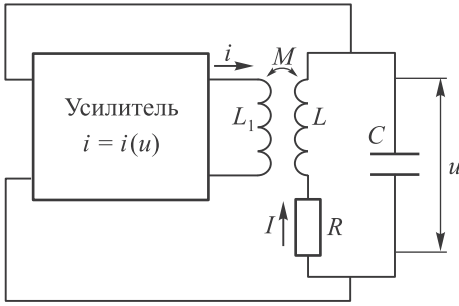


Рис. 26.7. Блок-схема автоколебательной системы



**Рис. 26.8.** Схема радиотехнического генератора электрических колебаний

Усилитель имеет нелинейную вольт-амперную характеристику, которую можно аппроксимировать многочленом  $i(u) = g_1 u - g_3 u^3$ , где  $g_i$  — постоянные коэффициенты. Коэффициент усиления должен быть достаточно велик, чтобы компенсировать потери энергии в контуре на джоулево тепло. Однако из-за нелинейности вольт-амперной характеристики усилителя коэффициент усиления падает с повышением напряжения в контуре, поэтому колебания не растут до бесконечности, а устанавливается стабильный режим автоколебаний.

В монографии<sup>2</sup> разобран принцип действия большого числа автоколебательных систем. При этом автор не использует математический аппарат, стремясь пояснить физическую сущность явления на основе энергетических соображений. Вот пример того, как с помощью самых общих рассуждений он рассказывает о принципе действия обратной связи. Рассмотрим график зависимости от амплитуды колебаний  $a$  количества энергии  $E_-$ , отдаваемой колебательной системой, и количество энергии  $E_+$ , получаемой от источника. Обе кривые должны расти, и если они расположены так, как показано на рис. 26.9, то точка пересечения кривых соответствует балансу энергий, который достигается при некоторой установившейся амплитуде  $a_0$ . Предположим теперь, что амплитуда колебаний по каким-то причинам слегка возросла и стала больше  $a_0$  (правая вертикальная пунктирная линия). Тогда поступление энергии станет меньше ее расхода ( $E_+ < E_-$ ), и колебания начнут затухать, их амплитуда станет уменьшаться. Если же, наоборот, амплитуда установившихся колебаний вдруг станет меньше стационарного значения  $a_0$  (левая пунктирная вертикальная линия), то поступление энергии превысит ее расход ( $E_+ > E_-$ ), что приведет к росту амплитуды. И в том, и в другом случае обратная связь препятствует отклонению от состояния динамического равновесия колебательной системы. Стало быть, такие колебания устойчивы. Энергетические кривые пересекаются также в начале координат, т. е. в состоянии покоя колебательной системы — в состоянии статического равновесия. Но это равновесие неустойчиво. При возникновении флуктуации, какого-то небольшого колебания, система начнет получать энергии больше, чем отдает, а это приведет к усилению флуктуаций и к росту амплитуды, пока

Примером автоколебательной системы может служить генератор электромагнитных колебаний, принципиальная схема которого показана на рис. 26.8 (взято из книги<sup>1</sup>). Источником энергии служит усилитель, включенный в электросеть. Наличие сопротивления  $R$  в колебательном контуре  $RLC$  обеспечивает необходимое затухание. Усилитель связан с колебательным контуром через катушку индуктивности  $L_1$ . Роль клапана выполняет электрод, на который подается напряжение  $u$  из колебательного контура.

<sup>1</sup> А. П. Кузнецов, С. П. Кузнецов, Н. М. Рыскин. Нелинейные колебания. М.: Физматлит, 2002.

<sup>2</sup> А. А. Харкевич. Автоколебания. М.: Гостехиздат, 1954.

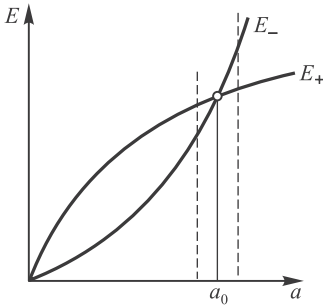


Рис. 26.9. К пояснению мягкого самовозбуждения автоколебаний

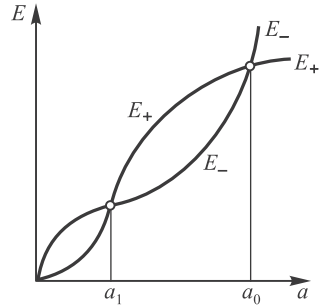
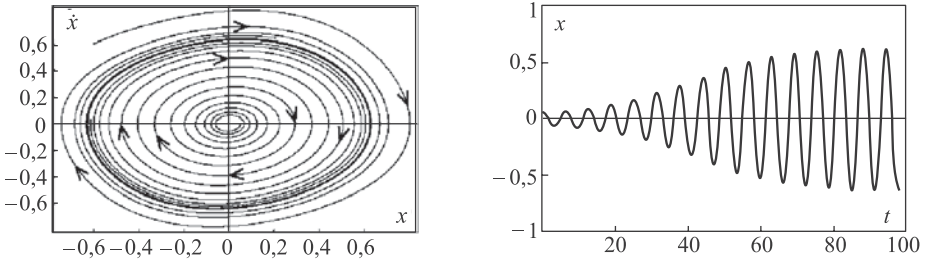


Рис. 26.10. К пояснению жесткого самовозбуждения автоколебаний

она не достигнет равновесного значения  $a_0$ . Этот процесс называется *мягким самовозбуждением* автоколебаний, такая система входит в автоколебательный режим самопроизвольно.

Возможно, однако, иное расположение энергетических кривых (рис. 26.10), где есть три точки пересечения кривых. Состояние статического равновесия (начало координат, колебаний нет) здесь устойчиво и самовозбуждения от флуктуаций здесь не происходит. Точка, соответствующая амплитуде  $a_1$ , — это состояние неустойчивого динамического равновесия, а колебание с амплитудой  $a_2$  устойчиво. Чтобы его возбудить, системе надо придать определенный толчок конечной величины (например, отклонить маятник остановившихся часов). Кривые  $E_+$  и  $E_-$  могут пересекаться и в большем числе точек, и все точки устойчивого равновесия на энергетическом графике — это разные автоколебания. Вся игра, как мы видим, строится на различной зависимости энергии от амплитуды. Для линейных систем зависимость была бы одинаковой — параболической, а две параболы с минимумом в начале координат других точек пересечения не имеют. Стало быть, автоколебания в линейной системе невозможны.

Этот раздел мы закончим фазовым портретом автоколебаний. Наличие установившихся колебаний, не зависящих от начальных условий, означает, что аттрактор в автоколебательной системе представляет собой не точку, а замкнутую кривую в фазовом пространстве — так называемый *предельный цикл*. Независимость автоколебаний от начальных условий означает, что все фазовые кривые независимо от того, где они начались, притягиваются к аттрактору, как бы наматываются на него. После того как в системе затухнут переходные процессы, установившимся автоколебаниям будет соответствовать движение по предельному циклу, аттрактору. Может случиться так, что в системе существует несколько аттракторов, и тогда у каждого из них будет существовать свой *бассейн* — область, откуда фазовые кривые стремятся именно к данному аттрактору, а не к какому-то другому. Это означает, что предельный цикл является изолированной замкнутой кривой — изолированной в том смысле, что в какой-то ее конечной окрестности других замкнутых кривых нет. Этим аттрактор автоколебательной системы отличается от замкнутых кривых нелинейного незатухающего осциллятора — там замкнутые фазовые траектории можно провести через любую точку фазового пространства.



**Рис. 26.11.** Для автоколебаний в радиотехническом генераторе показаны фазовый портрет (слева) и временная развертка (справа)

Сказанное иллюстрируется рисунком 26.11 (взято из книги<sup>1</sup>, см. сноску на с. 52). Колебания в радиотехническом генераторе (см. рис. 26.8) могут быть описаны моделью осциллятора Ван-дер-Поля:

$$\ddot{x} - (\lambda - x^2)\dot{x} + x = 0, \quad (26.12)$$

где координата  $x(t)$  соответствует напряжению на конденсаторе, а параметр  $\lambda$  характеризует силу обратной связи колебательного контура с усилителем. Режим самовозбуждения соответствует положительным значениям  $\lambda > 0$ . При слабой обратной связи (на рисунке взято значение  $\lambda = 0,1$ ) даже небольшая начальная флуктуация тока  $\dot{x}(0)$  ведет к нарастанию колебаний, причем требуется значительное число колебаний, прежде чем система войдет в установившийся автоколебательный режим (см. рис. 26.11, справа). Фазовый портрет колебаний показан на том же рисунке слева. Отчетливо видно, как колебания стремятся к предельному циклу.

## 26.4 Эволюция и взаимодействие популяций

В этом разделе мы познакомимся с описанием сообществ (популяций) с помощью простых математических моделей. Элементарными единицами популяций могут быть химические молекулы, клетки или организмы. В любом случае основные законы имеют общий характер и описываются одинаковыми уравнениями. Переменная величина, подчиняющаяся этим уравнениям, может быть полной массой или числом элементарных единиц популяции. Решения уравнений демонстрируют основные принципы динамического равновесия, реализуемого в природе и, возможно, в социальных структурах.

### Популяция в отсутствие сдерживающих факторов

Самым простым случаем является ничем не сдерживаемое развитие популяции. Пусть  $n(t)$  — число ее элементов как функция времени, а  $n_0 = n(0)$  — начальное состояние популяции. Основным параметром, характеризующим развитие популяции, является коэффициент прироста  $\varepsilon$ , определяемый как относительное превышение рождаемости над смертностью в единицу времени  $\varepsilon = dn/(ndt)$ . Примером такого коэффициента является, например, процент прироста населения за год. Обратная величина  $\tau = 1/\varepsilon$  имеет размерность времени и определяет

характерное время размножения популяции. Получающееся дифференциальное уравнение первого порядка

$$\frac{dn}{dt} = \frac{n}{\tau} \quad (26.13)$$

легко интегрируется:

$$\int_{n_0}^{n(t)} \frac{dn}{n} = \ln \frac{n(t)}{n_0} = \frac{1}{\tau} \int_0^t dt = \frac{t}{\tau}, \quad (26.14)$$

так что его решение можно записать в виде

$$n(t) = n_0 e^{t/\tau}. \quad (26.15)$$

Физический смысл времени  $\tau$  — это промежуток, за который численность популяции увеличивается в  $e \approx 2,718$  раза. Экспоненциальный рост численности популяций называется законом Мальтуса (1798) и представляет собой предельный случай наиболее благоприятного ее развития, которое не сдерживается никакими внешними факторами. Мы говорим о росте, так как обычно рождаемость превышает смертность и численность популяции увеличивается. Это соответствует положительному знаку  $\varepsilon$ . Если же  $\varepsilon < 0$ , то решение (26.15) описывает экспоненциальное убывание численности популяции:  $n(t) = n_0 \exp(-t/\tau)$ ,  $\tau = 1/|\varepsilon|$ . По такому закону, например, убывает число распадающихся атомов в радиоактивном образце.

## Внутривидовая конкуренция

Закон Мальтуса приводит к быстрому росту популяций. Если бы он был верным всегда и везде, то наш мир был бы переполнен, скажем, бактериями и микробами, а число начальников быстро превысило бы число управляемых ими людей. Как мы знаем, этого, к счастью, не происходит. Причиной тому — конкуренция между элементами популяций, внутривидовая борьба за общие средства к существованию, будь то зарплата, пища или любой другой жизненно важный продукт. В условиях нехватки этого продукта размножение популяции сдерживается столкновениями между ее элементами. Число столкновений пропорционально квадрату численности популяции (каждый конкурирует с каждым), а степень «жесткости» столкновений характеризуется коэффициентом  $h$ . Поэтому можно принять, что столкновения ведут к убыли численности популяции со скоростью  $-hn^2$ , так что вместо (26.13) мы получим уравнение  $\dot{n} = n/\tau - hn^2$ . Обратная коэффициенту  $h$  величина имеет размерность времени. Параметр с той же размерностью у нас уже введен — это время размножения популяции. Поэтому вместо  $h$  удобно ввести новую безразмерную величину  $n_* = 1/h\tau$ , физический смысл которой прояснится чуть позже. Тогда уравнение для эволюции популяции можно записать в виде

$$\frac{dn}{dt} = \frac{1}{\tau} \left( n - \frac{n^2}{n_*} \right). \quad (26.16)$$

Первый член в правой части соответствует свободному размножению популяции, а второй член с отрицательным знаком — убыли численности популяции вследствие конкуренции, столкновений между ее членами. Это уравнение называется уравнением Ферхюльста (1838).



Прежде всего обратим внимание, что уравнение (26.16) допускает стационарное, не зависящее от времени решение. Из условия стационарности  $dn/dt = 0$  следует равенство нулю правой части уравнения, откуда, кроме тривиального решения  $n = 0$  (отсутствие популяции вообще), следует и стационарное решение  $n = n_*$ . Таким образом проясняется физический смысл введенного параметра  $n_*$  — это стационарная (равновесная) численность популяции. Если начальная численность популяции  $n_0$  равна стационарной, то таковой она останется навсегда. Но что произойдет, если  $n_0 \neq n_*$ ?

Уравнение (26.16) с ненулевой правой частью также легко интегрируется:

$$\int_{n_0}^{n(t)} \frac{n_*}{n(n_* - n)} dn = \frac{t}{\tau}. \quad (26.17)$$

Преобразуя подынтегральное выражение к виду

$$\frac{n_*}{n(n_* - n)} = \frac{1}{n} + \frac{1}{n_* - n}, \quad (26.18)$$

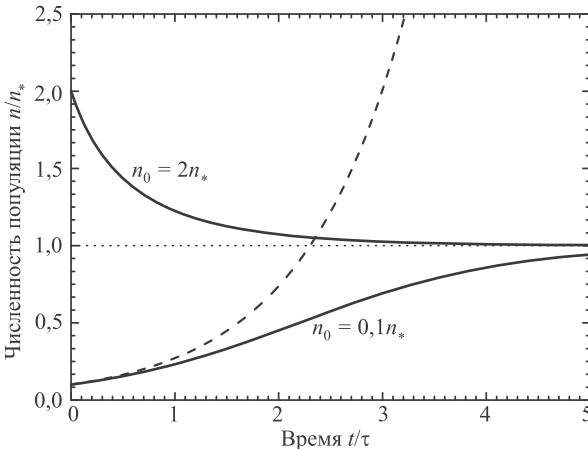
легко выполняем интегрирование и получаем

$$\frac{t}{\tau} = \ln \frac{n}{n_0} - \ln \frac{n_* - n}{n_* - n_0} = \ln \frac{n(n_* - n_0)}{n_0(n_* - n)}, \quad (26.19)$$

откуда следует окончательное решение

$$n(t) = \frac{n_0 n_*}{n_0 + (n_* - n_0)e^{-t/\tau}}. \quad (26.20)$$

В начальный момент времени  $t = 0$  экспонента равна единице и численность популяции действительно равна  $n_0$ . Далее численность популяции возрастает (если  $n_0 < n_*$ ) или убывает (если  $n_0 > n_*$ ), стремясь асимптотически к равновесному значению. На рисунке 26.12 в относительных единицах  $x = t/\tau$  и



**Рис. 26.12.** Эволюция популяции при двух разных начальных условиях:  $n_0 = 0,1 n_*$  и  $n_0 = 2n_*$ . Штриховая линия соответствует эволюции популяции при первом из указанных начальных условий при отсутствии внутривидовой конкуренции

$y = n/n_*$  показано изменение численности популяции для начальных значений  $n_0 = 0, 1n_*$  и  $n_0 = 2n_*$ . Видно, что за время порядка нескольких времен размножения популяции ее численность выходит на стационарное значение. На этом же графике штриховой линией показано решение уравнения (26.13) для первого из указанных начальных значений. Напомним, что этот случай соответствует отсутствию внутривидовой конкуренции ( $h = 0$ ) или, что то же самое, бесконечно большой равновесной численности популяции ( $n_* \rightarrow \infty$ ). В реальной жизни так не бывает, но из рисунка видно, что закон Мальтуса справедлив при не слишком больших временах эволюции, когда численность популяции не слишком велика и внутривидовая конкуренция не носит острого характера.

### О взаимоотношениях зайцев и волков

В этом разделе мы рассмотрим взаимодействие двух популяций, одна из которых (жертвы) служит пищей для другой (хищники). Уравнения известной модели Лотки—Вольтерры (1925–1926) имеют вид

$$\frac{dn}{dt} = \frac{1}{\tau_n} \left( n - \frac{nN}{N_*} \right), \quad (26.21)$$

$$\frac{dN}{dt} = \frac{1}{\tau_N} \left( -N + \frac{nN}{n_*} \right), \quad (26.22)$$

где функции времени  $n(t)$  и  $N(t)$  представляют собой численность популяций жертв и хищников, соответственно. В основе уравнений лежит самая упрощенная модель взаимоотношений жертвы и хищника.

Внутривидовая конкуренция в них не учитывается и предполагается, что в отсутствие врага жертвы будут размножаться неограниченно (т.е. питания им заведомо хватит). Такая ситуация описывается первым членом в правой части уравнения (26.21). Параметр  $\tau_n$  характеризует скорость размножения жертв в отсутствие хищников. Второй член, входящий в правую часть (26.21), описывает убыль количества жертв из-за их встреч с хищниками. Эта убыль пропорциональна числу встреч жертв с хищниками, которое, в свою очередь, пропорционально произведению  $nN$  численностей их популяций. Уравнение (26.22) описывает изменение численности популяции хищников. Предполагается, что в отсутствие жертв число хищников будет быстро уменьшаться из-за недостатка пищи (т.е. смертность превысит рождаемость). Поэтому первый член в правой части (26.22) имеет отрицательный знак. Скорость вымирания хищников численно характеризуется временем  $\tau_N$ , за которое их число уменьшилось бы в  $e$  раз. Однако встречи с жертвами позволяют хищникам выжить, что описано вторым, положительным, членом в правой части (26.22). Число таких встреч опять-таки пропорционально произведению  $nN$ .

Очевидно, что пара чисел  $n_*$  и  $N_*$  является стационарным решением этих уравнений. Ясно, что если начальная численность популяций случайно оказалась равной стационарным значениям, то таковой она и останется. Если же начальная численность была иной, то численность популяций будет периодически меняться со временем. В дальнейшем удобно выразить численность популяций в единицах

их стационарных значений, а время в единицах «времени вымирания» хищников  $\tau_N$ . Введем новые переменные:

$$\tau = t/\tau_N, \quad x(\tau) = \frac{n}{n_*}, \quad y(\tau) = \frac{N}{N_*}. \quad (26.23)$$

Стационарной численности популяций соответствуют значения  $x_* = y_* = 1$ . С учетом соотношений (26.23) уравнения (26.21), (26.22) можно записать в виде

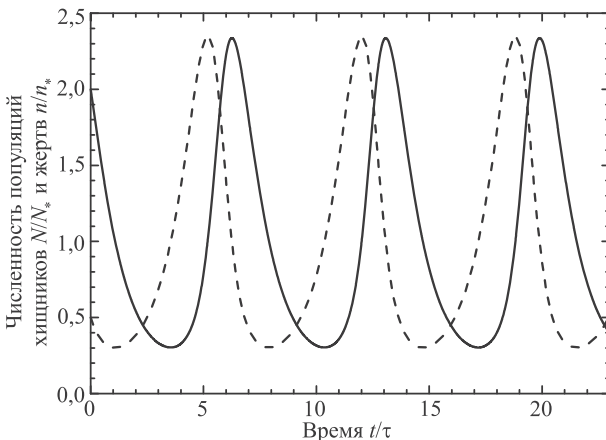
$$\begin{aligned} \frac{dx}{d\tau} &= qx(1-y), \\ \frac{dy}{d\tau} &= -y(1-x). \end{aligned} \quad (26.24)$$

Преимуществом введенных единиц является то, что теперь мы имеем дело лишь с одним безразмерным параметром задачи — величиной  $q$ , равной отношению характерных времен размножения жертв и вымирания хищников:

$$q = \tau_N/\tau_n. \quad (26.25)$$

На рисунке 26.13 показана зависимость от времени численности популяций жертв и хищников для начальных условий  $n_0 = 0,5n_*$  и  $N_0 = 2N_*$ , полученная в результате численного решения уравнений (26.24) в случае, когда характерные времена развития обеих популяций совпадают ( $q = 1$ ). Поскольку в начальный момент времени  $t = 0$  число хищников заметно превышает стационарное значение, они начинают вымирать. Число же жертв, еще «не знающих» об этом, также уменьшается, достигает минимума, после чего ситуация становится для жертв более благоприятной и они начинают размножаться. Хищники еще этого «не чувствуют» и продолжают вымирать. Это запаздывание воздействия одной популяции на другую проявляется в сдвиге по фазе периодических кривых на рис. 26.13.

Интересна также зависимость численности популяций не от времени, а друг от друга. Не выписанные нами решения  $x(\tau)$  и  $y(\tau)$  задают неявно искомую



**Рис. 26.13.** Зависимость от времени численности популяций хищников (сплошная линия) и жертв (штриховая линия) для начальных условий  $N_0 = 2N_*$ ,  $n_0 = 0,5n_*$

функцию  $y(x)$ . Учитывая, что

$$y'(x) \equiv \frac{dy}{dx} = \frac{dy/d\tau}{dx/d\tau} = \frac{\dot{y}}{\dot{x}},$$

делением второго из уравнений (26.24) на первое мы получаем дифференциальное уравнение первого порядка с разделяющимися переменными, из которого исключена явная зависимость от времени:

$$y'(x) = -\frac{1}{q} \frac{y}{1-y} \frac{1-x}{x}. \quad (26.26)$$

Уравнение (26.26) легко интегрируется:

$$q \int \frac{1-y}{y} dy = - \int \frac{1-x}{x} dx. \quad (26.27)$$

Из (26.27) немедленно получаем

$$q(\ln y - y) = -(\ln x - x) + \text{const}. \quad (26.28)$$

Постоянную интегрирования определяем из начальных условий  $x_0$ ,  $y_0$ , так что уравнение (26.28) может быть записано в виде

$$q \ln(y/y_0) - q(y - y_0) = -\ln(x/x_0) + (x - x_0). \quad (26.29)$$

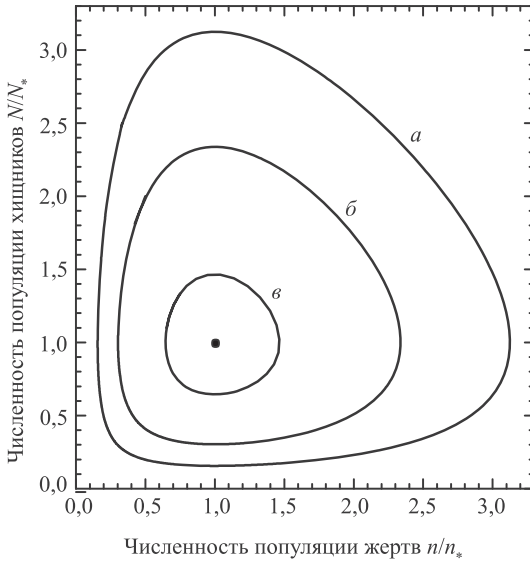
Потенцируя уравнение (26.29), получаем искомое соотношение между численностью популяций жертвы и хищника

$$\left(\frac{y}{y_0}\right)^q e^{-q(y-y_0)} = \frac{x_0}{x} e^{(x-x_0)}, \quad (26.30)$$

т. е. неявно заданную функцию  $y(x)$ .

Можно показать, что решение задачи существует при любых начальных значениях  $x_0$ ,  $y_0$ , причем каждому значению  $x$  соответствуют два значения  $y$  и, наоборот, каждому значению  $y$  соответствуют два значения  $x$ . Периодичность решений во времени означает, что график функции  $y(x)$  представляет собой замкнутую кривую, расположенную в первой четверти плоскости  $xy$ . Естественно, кривая проходит через точку с координатами  $(x_0, y_0)$ . Меняя начальные условия, мы получаем целое семейство таких кривых. На рисунке 26.14 для примера показаны три кривые, соответствующие начальным значениям  $x_0 = 1/3$ ,  $y_0 = 5/2$  (а),  $x_0 = 1/2$ ,  $y_0 = 2$  (б) и  $x_0 = 3/4$ ,  $y_0 = 4/3$  (в). По мере приближения начальных значений к стационарным, охватываемая кривыми область становится меньше и в предельном случае  $x_0 = 1$ ,  $y_0 = 1$  стягивается в точку, соответствующую стационарному решению (также показана на рис. 26.14).

Рассматривая рис. 26.14, можно заметить, что при равновесном значении численности жертв  $x = x_*$  численность хищников  $y$  достигает своих максимального и минимального значений. Аналогично численность популяции жертв  $x$  достигает своих минимального и максимального значений при стационарном значении численности популяции хищников  $y = y_*$ .



**Рис. 26.14.** Кривые взаимной зависимости численности популяций жертв и хищников для трех различных начальных условий. Точка в центре графика — стационарное состояние популяций

Приведенные примеры демонстрируют, что в природе практически невозможны стационарные состояния: они могут возникнуть разве что случайно. Нормой, скорее, являются периодические колебания вокруг стационарных значений, т. е. динамическое равновесие. Формально из наших результатов следует, что популяция зайцев не должна исчезнуть из-за аппетита волков — решения нигде не обращаются в ноль. Но они могут подходить сколь угодно близко к нулю, т. е. при слишком большом количестве хищников зайцам грозит полное уничтожение. Скажем, если в лесу остается пара зайцев, то дело случайности, дадут ли они достаточно потомства для нового витка периодического процесса.

Заметим в заключение, что уравнения Лотки—Вольтерры дают простейшее (и не всегда удовлетворительное) описание системы «жертва — хищник». Существуют более сложные обобщения этих уравнений, в том числе на три и более популяций. Система «жертва — хищник» рассматривается не только в биологии, но и в других областях науки (в частности, в экономике).

## Контрольные вопросы

1. При каких условиях даже малые колебания не будут гармоническими?
2. Что такое изохронные колебания?
3. Что такое фазовый портрет колебаний?
4. Что такое аттрактор?
5. Дайте определение автоколебаний.
6. Каковы характерные особенности автоколебаний?
7. Каковы характерные особенности фазового портрета автоколебаний?

## Глава 27

# Волновые процессы

В предыдущих главах изучались колебания в системах с одной, максимум с двумя степенями свободы. Теперь мы переходим к системам с бесконечным числом степеней свободы. Примерами могут служить колебания в газовой среде, колебания твердых тел, натянутой струны и т. п. Во всех этих примерах колебания происходят в каждой точке среды или тела, которую можно рассматривать как осциллятор. Соседние осцилляторы связаны между собой, так что между ними возможны процессы передачи энергии. В таких случаях говорят о распространяющейся волне.

### 27.1 Волны в упругих средах

#### Колебания струны

Рассмотрим малые колебания струны, натянутой силой  $F$  вдоль оси  $x$ . Произвольная точка струны с координатой  $x$  в момент времени  $t$  смещается на вектор  $\vec{u}(x, t)$ . Ограничимся простейшим колебательным процессом, когда все векторы смещения  $\vec{u}(x, t)$  в любой момент времени перпендикулярны оси  $x$  и лежат в одной плоскости. Тогда смещения точек струны можно описать одной функцией  $u(x, t)$ , характеризующей вертикальное перемещение струны (рис. 27.1).

Напряжения, возникающие в струне, направлены по касательным к ее мгновенному профилю. Мы будем рассматривать малые колебания, пренебрегая дополнительными силами упругости, возникающими из-за удлинения струны при движении ее элементов. Тогда натяжение струны можно считать постоянным для всех времен  $t$  и точек  $x$ . Выделим элемент струны, лежащий между координатами  $x$  и  $x + \Delta x$ . Рассмотрим точку  $x$ . Тангенс угла наклона силы  $F$ , действующей

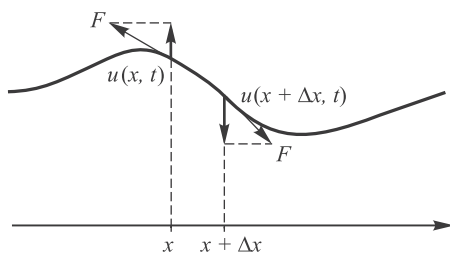


Рис. 27.1. Колебания струны

на этот край элемента, равен  $\operatorname{tg} \varphi = u'(x, t)$  (штрих означает производную по координате). Вертикальная компонента силы равна  $F_y(x, t) = F \sin \varphi$ . Так как угол  $\varphi$  мал,  $\operatorname{tg} \varphi \approx \varphi \approx \sin \varphi$ . Тогда  $F_y(x, t) = F u'(x, t)$ . Аналогично вертикальная компонента силы натяжения струны, действующей на другом конце выделенного элемента, равна  $F_y(x + \Delta x, t) = F u'(x + \Delta x, t)$ . Равнодействующая этих сил равна:

$$\Delta F = F (u'(x + \Delta x, t) - u'(x, t)) \approx F u''(x, t) \Delta x.$$

Заметим, что горизонтальные компоненты силы натяжения  $F_x = F \cos \varphi \approx F$  не зависят от положения точки и потому их равнодействующая равна нулю. Это означает, что элементы струны движутся только в вертикальном направлении.

Если линейная плотность (масса единицы длины) струны равна  $\rho$ , то масса элемента равна  $\Delta m = \rho \Delta x$ . Записываем уравнение второго закона Ньютона для вертикального смещения элемента струны:  $\Delta m \ddot{u}(x, t) = \Delta F$  (здесь точка означает производную по времени). Подставляя сюда выражение для  $\Delta F$ , получаем уравнение движения струны:

$$\rho \ddot{u}(x, t) = F u''(x, t). \quad (27.1)$$

Это уравнение можно переписать в виде

$$\frac{\partial^2}{\partial x^2} u(x, t) - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} u(x, t) = 0, \quad v = \sqrt{\frac{F}{\rho}}. \quad (27.2)$$

Определим размерность величины  $v$ . Размерность силы  $[F] = \text{MLT}^{-2}$ , размерность линейной плотности  $[\rho] = \text{ML}^{-1}$ . Отсюда

$$[v] = \sqrt{\frac{\text{MLT}^{-2}}{\text{ML}^{-1}}} = \text{LT}^{-1},$$

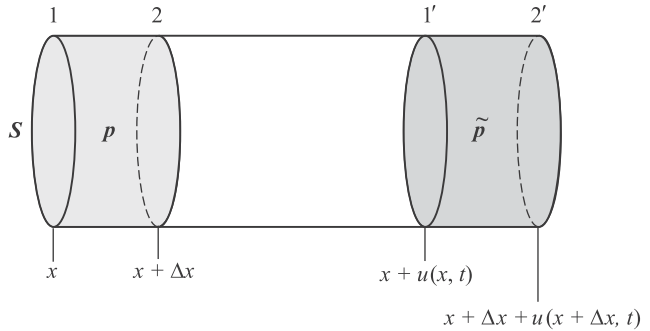
т. е. величина  $v$  имеет размерность скорости.

## Колебания в идеальном газе

Рассмотрим колебания в газе, происходящие вдоль одной оси  $x$ . В отличие от струны частицы газа смещаются здесь в продольном направлении, и величины смещения мы будем обозначать тем же символом  $u(x, t)$ .

Рассмотрим элементарный цилиндрический объем газа длиной  $\Delta x$ , ограниченный сечениями 1 и 2 площадью  $S$ , находящимися в точке с координатами  $x$  и  $x + \Delta x$  (рис. 27.2). Масса газа в этом объеме равна  $\Delta m = \rho S \Delta x$ , где  $\rho$  — плотность газа.

Пусть в стационарном состоянии давление газа равно  $p$ . При колебаниях объем смещается в новое положение между сечениями 1' и 2' с координатами  $x + u(x, t)$  и  $x + \Delta x + u(x + \Delta x, t) \approx x + u(x, t) + \Delta x [1 + u'(x, t)]$ . Объем газа в новом положении становится равным  $S \Delta x [1 + u'(x, t)]$ , а давление в нем —  $\tilde{p}$ . Найдем это давление. Колебательные процессы в газах происходят достаточно быстро, так что можно считать, что элементарный объем не успевает обмениваться теплотой



**Рис. 27.2.** Колебания в газе

с соседними объемами. Значит, процесс можно считать адиабатным. Записываем уравнение этого процесса (см. т. 1, разд. 11.4):

$$p (S\Delta x)^\gamma = \tilde{p} (S\Delta x [1 + u'(x, t)])^\gamma. \quad (27.3)$$

Отсюда

$$\tilde{p} = \frac{p}{[1 + u'(x, t)]^\gamma} \approx p[1 - \gamma u'(x, t)], \quad (27.4)$$

где  $\gamma$  — показатель адиабаты, зависящий от вида газа. Мы использовали также малость смещения  $u$ , записав

$$\frac{1}{[1 + u'(x, t)]^\gamma} \approx \frac{1}{1 + \gamma u'(x, t)} \approx 1 - \gamma u'(x, t).$$

Составим теперь уравнение движения элементарного объема. Его ускорение равно  $\ddot{u}(x, t)$ . Сила, действующая на объем, определяется разностью давлений в сечениях 1' и 2':

$$\Delta F = S(\tilde{p}_{1'} - \tilde{p}_{2'}) \approx -\tilde{p}' S\Delta x. \quad (27.5)$$

Подставляя сюда выражение для давления  $\tilde{p}$ , находим:

$$\Delta F = \gamma p u''(x, t) S\Delta x. \quad (27.6)$$

Используем теперь (27.6) в уравнении второго закона Ньютона  $\Delta m \ddot{u} = \Delta F$ :

$$\rho S\Delta x \ddot{u}(x, t) = \gamma p u''(x, t) S\Delta x. \quad (27.7)$$

Это уравнение можно представить в виде

$$\frac{\partial^2}{\partial x^2} u(x, t) - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} u(x, t) = 0, \quad v = \sqrt{\gamma \frac{p}{\rho}}. \quad (27.8)$$

Величина  $v$  имеет размерность скорости. Уравнение колебаний газа совпало с уравнением колебаний струны, хотя они описывают процессы в совершенно различных физических системах.



## Колебания в твердых телах

Колебательные процессы в твердых телах похожи на колебания в газах. Пусть рисунок 27.2 представляет теперь деформацию твердого тела. Относительная деформация элементарного объема при смещении  $u$  равна:

$$\sigma = \frac{\Delta u}{\Delta x} = u'(x, t). \quad (27.9)$$

Согласно закону Гука, это приводит к появлению упругой силы:

$$F = SE\sigma = SEu'(x, t), \quad (27.10)$$

где  $E$  — коэффициент (*модуль Юнга*), характеризующий жесткость среды. Равнодействующая сил упругости, действующих в сечениях  $1'$  и  $2'$ , равна:

$$\Delta F = F_{2'} - F_{1'} = SEu''(x, t)\Delta x. \quad (27.11)$$

Записывая второй закон Ньютона

$$\rho S\Delta x \ddot{u}(x, t) = SEu''(x, t)\Delta x, \quad (27.12)$$

находим уравнение колебаний в твердом теле:

$$\frac{\partial^2}{\partial x^2} u(x, t) - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} u(x, t) = 0, \quad v = \sqrt{\frac{E}{\rho}}. \quad (27.13)$$

Размерность модуля Юнга совпадает с размерностью давления, так что величина  $v$  опять-таки имеет размерность скорости.

Мы рассматривали продольные смещения в твердом теле. В отличие от газов, упругие силы возникают в твердых телах и при деформации сдвига. Уравнение для таких поперечных колебаний имеет тот же вид (27.13), но вместо модуля Юнга в выражении для  $v$  будет стоять так называемый *модуль сдвига*  $G$ :

$$v = \sqrt{\frac{G}{\rho}}. \quad (27.14)$$

## 27.2 Решение волнового уравнения

Уравнение, описывающее колебания различных упругих сред, называется *волновым уравнением*. Запишем его формально в виде

$$\begin{aligned} & \frac{\partial^2}{\partial x^2} u(x, t) - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} u(x, t) = \\ & = \left( \frac{\partial}{\partial x} - \frac{1}{v} \frac{\partial}{\partial t} \right) \left( \frac{\partial}{\partial x} + \frac{1}{v} \frac{\partial}{\partial t} \right) u(x, t) = 0. \end{aligned} \quad (27.15)$$

Введем теперь вместо  $x, t$  новые переменные:

$$\begin{aligned} \xi &= x - vt, & \chi &= x + vt, \\ x &= \frac{\xi + \chi}{2}, & t &= \frac{\chi - \xi}{2v}. \end{aligned} \quad (27.16)$$

Производные по новым переменным выражаются по стандартным правилам дифференцирования сложной функции:

$$\begin{aligned}\frac{\partial}{\partial \xi} &= \frac{\partial x}{\partial \xi} \frac{\partial}{\partial x} + \frac{\partial t}{\partial \xi} \frac{\partial}{\partial t} = \frac{1}{2} \left( \frac{\partial}{\partial x} - \frac{1}{v} \frac{\partial}{\partial t} \right), \\ \frac{\partial}{\partial \chi} &= \frac{\partial x}{\partial \chi} \frac{\partial}{\partial x} + \frac{\partial t}{\partial \chi} \frac{\partial}{\partial t} = \frac{1}{2} \left( \frac{\partial}{\partial x} + \frac{1}{v} \frac{\partial}{\partial t} \right).\end{aligned}\quad (27.17)$$

Отсюда следует, что уравнение (27.15) в новых переменных можно записать в виде:

$$\frac{\partial}{\partial \chi} \frac{\partial}{\partial \xi} u = 0. \quad (27.18)$$

Так как производная по  $\chi$  равна нулю, то  $\partial u / \partial \xi$  не зависит от этой переменной и, стало быть, является некой функцией  $w$  только от переменной  $\xi$ :

$$\frac{\partial}{\partial \xi} u = w(\xi). \quad (27.19)$$

Интегрируем теперь это уравнение:

$$u = \int w(\xi) d\xi + f_2. \quad (27.20)$$

Первое слагаемое в правой части является только функцией переменной  $\xi$ , которую мы обозначим как  $f_1(\xi)$ . Второе слагаемое — постоянная интегрирования. Она не зависит от  $\xi$ , являясь, стало быть, функцией только переменной  $\chi$ :  $f_2 = f_2(\chi)$ .

Мы получили, что решение волнового уравнения имеет вид  $u = f_1(\xi) + f_2(\chi)$ . Подставляя сюда выражения (27.16), мы возвращаемся к переменным  $x, t$ :

$$u(x, t) = f_1(x - vt) + f_2(x + vt). \quad (27.21)$$

Функции  $f_1$  и  $f_2$  совершенно произвольны и должны быть определены из начальных условий.

Обсудим физический смысл полученных решений. Ограничимся сначала первым слагаемым. Пусть  $u(x, t) = f_1(x - vt)$ . В момент времени  $t = 0$  функция  $f_1(x)$  задает начальное распределение смещений (профиль струны, деформацию твердого тела, распределение давления или частиц в газе):  $u(x, 0) = f_1(x)$ . Предположим, например, что это распределение имеет максимум в точке  $x = x_0$  (рис. 27.3). Подобное распределение называют обычно *волновым пакетом*. В момент времени  $t$  волновой пакет не изменит свою форму, но сдвинется вправо на расстояние  $vt$ , так что максимум теперь придется на точку  $x = x_0 + vt$ . Мы видим, что начальное распределение движется вправо со скоростью  $v$ . Аналогично второе слагаемое  $f_2(x + vt)$  описывает движение волнового пакета влево с той же скоростью. Общее решение (27.21) является суперпозицией двух описанных.

В свою очередь, любой волновой пакет может быть представлен как суперпозиция гармонических функций. Отсюда особая роль решений волнового уравнения вида

$$u(x, t) = A \cos(\omega t - kx + \varphi_0). \quad (27.22)$$

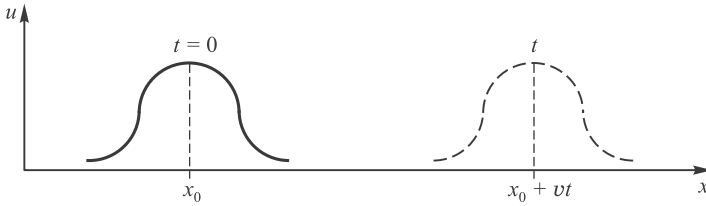


Рис. 27.3. Движение волнового пакета  $f_1(x - vt)$

Это решение описывает *монохроматическую волну*, распространяющуюся направо со скоростью

$$v = \frac{\omega}{k}. \quad (27.23)$$

Действительно, выражение (27.22) можно представить в виде

$$u(x, t) = A \cos k(x - vt - \varphi_0/k),$$

что является одной из бесчисленных возможностей конкретного воплощения функции  $f_1(x - vt)$  в (27.21). Величина  $\omega$  — это циклическая частота колебаний, а  $k$  — *волновое число* (модуль *волнового вектора* — см. с. 81).

Пусть наблюдатель находится в точке  $x = x_0$  и следит за колебаниями среды в этой точке. Он обнаружит, что колебательное движение происходит по закону

$$u(x_0, t) = A \cos(\omega t + \varphi), \quad \varphi = \varphi_0 - kx. \quad (27.24)$$

Наблюдатель в другой точке также обнаружит гармонические колебания с той же частотой, но с другой начальной фазой  $\varphi$ . Чем правее точка наблюдения, тем большее запаздывание по фазе имеют там колебания. Соответственно выражение

$$u(x, t) = A \cos(\omega t + kx + \varphi_0)$$

описывает монохроматическую волну, распространяющуюся налево.

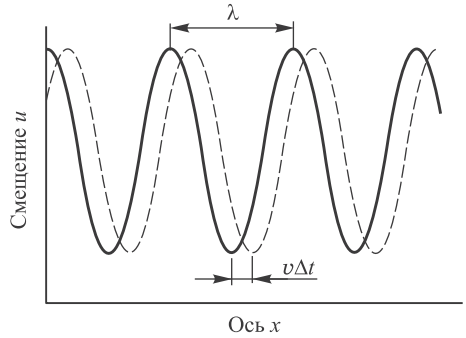
Проведем теперь другой мысленный опыт: «сфотографируем» нашу волну в какой-то момент времени  $t = t_0$  (в случае колеблющейся струны для этого даже не нужно изощренных приборов). На снимке мы увидим периодическую пространственную структуру:

$$u(x, t_0) = A \cos(\psi_0 - kx), \quad \psi_0 = \varphi_0 + \omega t_0. \quad (27.25)$$

Максимумы смещений приходятся на точки с координатами  $x_n$ :  $\psi_0 - kx_n = 2\pi n$  (рис. 27.4). Период повторения  $\lambda$  тех же смещений в пространстве есть расстояние между двумя ближайшими максимумами:  $\lambda = x_{n+1} - x_n$ . Получаем в итоге

$$\lambda = \frac{2\pi}{k}. \quad (27.26)$$

Величина  $\lambda$  называется *длиной волны*.



**Рис. 27.4.** Смещение точек среды в моменты времени  $t$  (сплошная кривая) и  $t + \Delta t$  (штриховая кривая)

Если «сфотографировать» волну в близкий к  $t_0$  момент времени  $t = t_0 + \Delta t$ , то на снимке вся пространственная структура сдвинется как целое на расстояние  $\Delta x = v\Delta t$ . Скорость  $v$  называется **фазовой скоростью волны**, так как с такой скоростью движутся максимумы, минимумы и вообще все точки с данным значением фазы. Используя уравнения (27.23) и (27.26), находим связь между характеристиками волны:

$$\lambda = v \frac{2\pi}{\omega} = \frac{v}{\nu}, \quad \text{или} \quad \lambda\nu = v. \quad (27.27)$$

Здесь  $\nu = \omega/2\pi$  — частота колебаний в волне.

Приведем численные примеры. Волна сгущений и разрежений в газе есть не что иное, как звуковая волна. Используя уравнение газового состояния, можно записать выражение для скорости (27.8) в виде

$$v = \sqrt{\gamma \frac{RT}{M}} = \sqrt{\gamma \frac{k_B T}{m}}, \quad (27.28)$$

где  $M$  — молярная масса газа,  $m$  — масса молекул, а  $T$  — абсолютная температура газа. Но среднеквадратическая скорость определяется из закона равнораспределения  $mv_{\text{тепл}}^2/2 = 3k_B T/2$ , откуда

$$v_{\text{тепл}} = \sqrt{\frac{3k_B T}{m}} = \sqrt{\frac{3}{\gamma}} v. \quad (27.29)$$

Иными словами, скорость звука в газе по порядку величины совпадает со скоростью теплового движения молекул, будучи меньше нее примерно в полтора раза ( $\sqrt{3/\gamma} \approx 1,5$ ). Молярная масса воздуха  $M = 29 \cdot 10^{-3}$  кг/моль, показатель адиабаты  $\gamma = 1,4$ . Подставляя эти значения в (27.28), находим скорость звука при комнатной температуре ( $T = 20^\circ\text{C} = 293$  К):

$$v = \sqrt{\frac{1,4 \times 8,31 \times 293}{29 \cdot 10^{-3}}} \approx 343 \text{ м/с}. \quad (27.30)$$

Человеческое ухо воспринимает частоты в диапазоне  $\nu = 20$  Гц — 20 кГц. Соответствующие длины волн равны  $\lambda = v/\nu = 343/20 \approx 17$  м для низких частот и  $\lambda = 343/20\,000 \approx 17$  мм для высоких.

Для стали модуль Юнга равен  $E = 20,6 \cdot 10^{10}$  Н/м<sup>2</sup>, модуль сдвига  $G = 8 \cdot 10^{10}$  Н/м<sup>2</sup>, а плотность  $\rho = 7,8 \cdot 10^3$  кг/м<sup>3</sup>. Соответственно получаем из (27.13), (27.14) скорости распространения продольных и поперечных колебаний:

$$\begin{aligned} v_{\parallel} &= \sqrt{\frac{E}{\rho}} = \sqrt{\frac{20,6 \cdot 10^{10}}{7,8 \cdot 10^3}} = 5,1 \cdot 10^3 \text{ м/с} = 5,1 \text{ км/с}, \\ v_{\perp} &= \sqrt{\frac{G}{\rho}} = \sqrt{\frac{8 \cdot 10^{10}}{7,8 \cdot 10^3}} = 3,2 \cdot 10^3 \text{ м/с} = 3,2 \text{ км/с}. \end{aligned} \quad (27.31)$$

Наконец, для воды роль модуля Юнга выполняет величина, обратная сжимаемости  $\kappa = 0,47 \cdot 10^{-9}$  Па<sup>-1</sup>. Плотность воды  $\rho = 10^3$  кг/м<sup>3</sup>. Для скорости звука в воде получаем:

$$v = \frac{1}{\sqrt{\kappa\rho}} = \frac{1}{\sqrt{0,47 \cdot 10^{-9} \times 10^3}} = 1,46 \cdot 10^3 \text{ м/с} = 1,46 \text{ км/с}. \quad (27.32)$$

Звук той же частоты будет иметь в воде и воздухе разные длины волн. Так, для  $\nu = 20$  кГц получаем длину волны в воде

$$\lambda = v/\nu = 1,46 \cdot 10^3 / 20 \cdot 10^3 = 0,073 \text{ м} = 7,3 \text{ см},$$

что надо сравнить с  $\lambda = 17$  мм в воздухе.

**Задача 27.2.** Для диагностики опухолей в мягких тканях применяется ультразвук с частотой  $\nu = 2$  МГц. Определить длину ультразвуковой волны в воздухе и в мягких тканях, где скорость распространения звука равна  $v = 1,5$  км/с.

**Решение.** Длина волны  $\lambda = v/\nu$ . В воздухе  $\lambda = 343/(2 \cdot 10^6) = 1,7 \cdot 10^{-4}$  м = 0,17 мм. В мягких тканях  $\lambda = 1,5 \cdot 10^3/(2 \cdot 10^6) = 7,5 \cdot 10^{-4}$  м = 0,75 мм. Как мы увидим в дальнейшем, длина волны любого излучения накладывает естественный предел на размеры объектов, которые можно различить с его помощью. Данный пример показывает, что диагностика опухолей, размеры которых меньше миллиметра, с помощью ультразвука затруднительна. ■

**Задача 27.3.** Летучая мышь использует для ориентирования ультразвук с частотой  $\nu = 100$  кГц. Определить размеры препятствий, которые заведомо не будут замечены летучей мышью. Ответить на тот же вопрос в отношении дельфинов, которые также используют эти частоты.

**Решение.** Длина волны, испускаемой летучей мышью, равна  $\lambda = v/\nu = 343/10^5 = 3,4 \cdot 10^{-3}$  м = 3,4 мм. Препятствия меньших размеров заведомо не могут быть замечены мышью с помощью испускаемой ультразвуковой волны. Для дельфинов ответ иной из-за другой скорости распространения звука в воде:  $\lambda = 1,46 \cdot 10^3/10^5 = 1,46 \cdot 10^{-2}$  м  $\approx 1,5$  см. Таким образом, летучая мышь может обнаружить насекомых, а дельфин — небольших рыбок. ■

**Задача 27.4.** Альпинист, спускающийся с отвесной скалы, висит на веревке длиной  $l = 30$  м. Страхующий его партнер подает ему сигнал, дергая веревку. За какое время  $t$  сигнал достигнет альпиниста? Масса альпиниста  $m = 80$  кг, масса одного метра веревки равна 75 г.

**Решение.** Так как можно найти линейную плотность веревки  $\rho = m/l = 7,5 \cdot 10^{-2}$  кг/м и силу ее натяжения  $F = mg$ , по формуле (27.2) находим скорость распространения колебаний:

$$v = \sqrt{\frac{mg}{\rho}} = \sqrt{\frac{80 \times 9,8}{7,5 \cdot 10^{-2}}} = 102 \text{ м/с.}$$

Отсюда определяем время прохождения сигнала:  $t = l/v = 30/102 \approx 0,3$  с. ■

## 27.3 Энергия волны

Рассмотрим для примера звуковую волну. Элемент объема на рис. 27.2 имеет кинетическую энергию

$$W_{\text{к}} = \frac{\Delta m \dot{u}^2}{2} = S \Delta x \frac{\rho \dot{u}^2}{2}. \quad (27.33)$$

При его деформации в данном объеме газа запасается потенциальная энергия  $W_{\text{п}}$ . Рассматривая колебания поршня, мы получили выражение (24.13) для силы упругости при перемещении поршня на расстояние  $x$ :

$$F = -\gamma \frac{p_0 S^2}{V_0} x. \quad (27.34)$$

Этот закон аналогичен закону Гука для силы упругости при сжатии или растяжении пружины. Следовательно, потенциальная энергия газа равна в данном случае

$$W_{\text{п}} = - \int_0^x F dx = \gamma \frac{p_0 S^2}{V_0} \frac{x^2}{2}. \quad (27.35)$$

Произведение  $Sx = \Delta V$  равно изменению объема газа под поршнем. Поэтому (27.35) можно записать в виде:

$$W_{\text{п}} = \gamma \frac{p_0}{V_0} \frac{\Delta V^2}{2}. \quad (27.36)$$

Применим это выражение к объему газа в звуковой волне. Давление в стационарном состоянии мы обозначили через  $p$ , следовательно, надо заменить  $p_0 \rightarrow p$ . Объем в стационарном состоянии равен  $V_0 = S \Delta x$ . Изменение объема при колебаниях равно:

$$\Delta V = S (u(x + \Delta x, t) - u(x, t)) \approx S \Delta x u'(x, t).$$

Получаем тогда для потенциальной энергии данного объема газа

$$W_{\text{п}} = \gamma \frac{p}{S \Delta x} \frac{(S \Delta x u')^2}{2} = S \Delta x \frac{\gamma p u'^2}{2}. \quad (27.37)$$

Сумма кинетической и потенциальной энергий равна полной энергии данного объема. Плотность энергии  $w$  в волне получаем, разделив полную энергию на величину объема:

$$w = \frac{W_{\text{к}} + W_{\text{п}}}{S \Delta x} = \frac{\rho \dot{u}^2}{2} + \frac{\gamma p u'^2}{2}. \quad (27.38)$$

Учитывая, что фазовая скорость волны равна  $v = \sqrt{\gamma p / \rho}$ , записываем (27.38) в виде

$$w = \frac{1}{2} \rho \left[ \left( \frac{\partial u(x, t)}{\partial t} \right)^2 + v^2 \left( \frac{\partial u(x, t)}{\partial x} \right)^2 \right]. \quad (27.39)$$

Точно такое же выражение получается для волны в твердом теле (неважно, продольной ли, поперечной ли) и для волны вдоль струны.

Подставляя сюда решение уравнения (27.22)  $u(x, t) = A \cos(\omega t - kx + \varphi_0)$  для монохроматической волны и учитывая соотношение  $v = \omega/k$ , получаем

$$w = \rho \omega^2 A^2 \sin^2(\omega t - kx + \varphi_0). \quad (27.40)$$

Плотность энергии волны различна в разных точках пространства и в разные моменты времени. Зафиксируем какую-то точку  $x$  и усредним плотность энергии в данной точке по времени. Среднее значение квадрата синуса равно  $1/2$ . Получаем тогда, что **среднее значение плотности энергии постоянно для всех точек среды и равно**:

$$\langle w \rangle = \frac{1}{2} \rho \omega^2 A^2. \quad (27.41)$$

Таким образом, среда обладает полным запасом энергии (кинетической и потенциальной), плотность которой пропорциональна плотности среды, квадрату циклической частоты и квадрату амплитуды. (Напомним, что для колебательной системы с одной степенью свободы энергия колебаний также пропорциональна квадрату частоты и квадрату амплитуды колебаний (см. (24.44).)

Если вернуться к выражению (27.40) для мгновенного значения плотности энергии, то легко убедиться, что любое взятое значение плотности энергии (например, ее максимум  $w_{\max} = \rho \omega^2 A^2 = 2\langle w \rangle$ ) перемещается вдоль оси  $x$  со скоростью  $v$ . Иными словами, волна переносит энергию. Эта энергия доставляется, естественно, от источника колебаний. Можно ввести также вектор **плотности потока энергии**  $\vec{j} = w\vec{v}$ , который равен энергии, переносимой волной в единицу времени через единичную площадку, ортогональную направлению распространения волны.

## Применения к звуковой волне

Напомним, что с точки зрения физики **звук** — это колебательное движение частиц упругой среды, распространяющееся в виде волн в газообразной, жидкой или твердой средах. Смещение частиц газа описывается стандартным выражением:

$$u(x, t) = A \cos(\omega t - kx + \varphi_0),$$

где  $A$  — амплитуда смещения, а  $v = \sqrt{\gamma p / \rho}$  — фазовая скорость волны. Под  $p$  и  $\rho$  понимаются давление и плотность невозмущенного газа. Смещение частиц

приводит к появлению избыточного давления  $\Delta p = \tilde{p} - p$ :

$$\begin{aligned}\Delta p &= -\gamma p \frac{\partial u(x, t)}{\partial x} = -\gamma p k A \sin(\omega t - kx + \varphi_0) = \\ &= -\frac{\gamma p \omega}{v} A \sin(\omega t - kx + \varphi_0).\end{aligned}\quad (27.42)$$

Здесь мы использовали соотношение (27.4). Учитывая, что  $\gamma p = \rho v^2$ , уравнение (27.42) можно переписать в виде

$$\Delta p = -\rho \omega v A \sin(\omega t - kx + \varphi_0).\quad (27.43)$$

Заметим, что колебания давления сдвинуты на  $\pi/2$  по отношению к колебаниям смещения частиц газа. При максимальном смещении  $u(x, t)_{\max} = A$  давление равно стационарному значению, следовательно,  $\Delta p = 0$ . Наоборот, амплитуда давления достигает максимума при нулевом смещении частиц газа.

**Интенсивностью**  $I$  волны называется среднее значение модуля плотности потока энергии, переносимой волной:

$$I = \langle w \rangle v = \frac{1}{2} \rho \omega^2 v A^2.\quad (27.44)$$

Звуковые волны принято характеризовать **уровнем громкости**  $L$ , измеряемым в децибелах (дБ). Связь уровня громкости с интенсивностью звуковой волны дается формулой:

$$L = 10 \lg \frac{I}{I_0}, \quad I_0 = 10^{-12} \text{ Вт/м}^2.\quad (27.45)$$

Выбор величины  $I_0$  связан с порогом слышимости в области частот 1000–4000 Гц, к которым наиболее восприимчиво ухо человека. Таким образом, при  $I = I_0$  уровень громкости полагается равным нулю. При интенсивностях волны порядка 1–10 Вт/м<sup>2</sup> волна перестает восприниматься как звук, вызывая только ощущение боли. Этому соответствует уровень громкости  $L = 120$ –130 дБ.

Найдем связь между интенсивностью звуковой волны, избыточным давлением  $\Delta p$ , создаваемым ею, и смещениями частиц газа. Амплитуда колебаний давления в волне равна (см. (27.43)):

$$\Delta p_{\max} = \rho \omega v A,$$

откуда

$$A = \frac{\Delta p_{\max}}{\rho \omega v}.\quad (27.46)$$

Интенсивность волны выражаем также через амплитуду давления:

$$I = \frac{1}{2} \rho \omega^2 v A^2 = \frac{\Delta p_{\max}^2}{2 \rho v}.\quad (27.47)$$

Отсюда находим избыточное давление в звуковой волне:

$$\Delta p_{\max} = \sqrt{2 \rho v I}.\quad (27.48)$$



Таблица 27.1. Уровни громкости некоторых звуков

Звук	$L$ , дБ
Шелест листьев	10
Шепот, тиканье часов	20
Уличный шум (без автотранспорта)	30
Нормальный разговор	60
Крик	80
Рок-группа	110
Болевой предел	120
Реактивный двигатель на расстоянии 50 м	130
Стартующая ракета на расстоянии 50 м	200

Учитывая, что плотность воздуха при нормальных условиях равна  $\rho = 1,2 \text{ кг/м}^3$ , получаем амплитуду колебаний давления на болевом пороге ( $L = 120 \text{ дБ}$ ,  $I = 1 \text{ Вт/м}^2$ ):

$$\Delta p_{\max} = \sqrt{2 \times 1,2 \times 343 \times 1} \approx 29 \text{ Па.}$$

Амплитуда смещения частиц газа зависит при этом от частоты:

$$A = \frac{\Delta p_{\max}}{\rho \omega v} = \frac{29}{1,2 \times 343 \times 2\pi\nu} = \frac{1,12 \cdot 10^{-2} \text{ м}}{\nu/1 \text{ Гц}}.$$

Отсюда следует, что при громкости  $L = 120 \text{ дБ}$  и частоте  $\nu = 20 \text{ Гц}$  смещение составляет  $A = 5,6 \cdot 10^{-4} \text{ м} = 0,56 \text{ мм}$ , а на частоте  $\nu = 20 \text{ кГц}$  смещение равно  $A = 5,6 \cdot 10^{-7} \text{ м} = 0,56 \text{ мкм}$ .

Найдем теперь амплитуду колебаний скорости частиц газа:

$$\dot{i}_{\max} = \omega A = \frac{\Delta p_{\max}}{\rho v}. \quad (27.49)$$

Она не зависит от частоты волны и при громкости  $L = 120 \text{ дБ}$  равна:

$$\dot{i}_{\max} = \frac{29}{1,2 \times 343} = 0,070 \text{ м/с} = 7,0 \text{ см/с}.$$

В таблице 27.1 приведены значения уровня громкости для некоторых звуков, с помощью которых можно найти избыточное давление, смещения и скорости частиц газа в некоторых случаях.

## 27.4 Стоячие волны

Помимо бегущих волн, рассмотренных в предыдущем разделе, в природе существуют и так называемые **стоячие волны**, образующиеся в результате специального случая суперпозиции (сложения) бегущих волн. Мы постоянно встречаемся с ними в своей практической жизни: когда говорим, поем, слушаем музыку. В этом разделе после общих математических формул мы совсем коротко обсуждаем некоторые вопросы музыкальной акустики в надежде, что эта часть не покажется нашим студентам самой скучной.

### Струна, закрепленная на одном конце

Предположим, что струна закреплена неподвижно в точке с координатой  $x = 0$  и натянута в отрицательном направлении оси  $x$ . Пусть по струне слева направо (т. е. в положительном направлении оси  $x$ ) распространяется волна

$$u_{\text{пад}} = A \cos(\omega t - kx). \quad (27.50)$$

Дойдя до точки закрепления, волна отразится. Если пренебречь потерями энергии, то амплитуда отраженной волны совпадет с амплитудой падающей волны. Надо учесть также, что при отражении происходит изменение направления движения элемента на обратное (как при упругом ударе шарика о стенку):

$$u_{\text{отр}} = -A \cos(\omega t + kx). \quad (27.51)$$

Суперпозиция падающей и отраженной волн имеет вид

$$u(x, t) = u_{\text{пад}} + u_{\text{отр}} = A (\cos(\omega t - kx) - \cos(\omega t + kx)). \quad (27.52)$$

Мы видим, что в любой момент времени  $u(0, t) \equiv 0$  — это и есть условие закрепления струны в точке  $x = 0$ . Воспользовавшись формулами тригонометрии для преобразования разности косинусов, запишем (27.52) в виде

$$u(x, t) = u_{\text{max}} \sin kx \sin \omega t, \quad (27.53)$$

где  $u_{\text{max}} = 2A$  — амплитуда стоячей волны.

Мы нашли особый тип колебаний: каждая точка струны колеблется с частотой  $\omega$ , причем все точки струны одновременно достигают своих максимальных отклонений  $u_{\text{max}} \sin kx$  (или проходят положение равновесия). Если мы находимся, например, в точке с координатой  $x_n = \pi n/k$ , то в любой момент времени смещение этой точки равно нулю, т. е. она остается неподвижной. Эти неподвижные точки называются **узлами** стоячей волны. Точки, где колебания струны достигают максимального значения, также не меняют своих положений вдоль струны: эти элементы струны колеблются в поперечном направлении, но не смещаются в продольном. То же самое относится и ко всем другим элементам струны. Такие колебания называются **стоячими волнами**. (Заметим, впрочем, что в стоячих волнах в газообразных и жидких средах колебания совершаются в продольном направлении. Таковы, например, колебания воздуха в духовых музыкальных инструментах. Здесь смещения не испытывают средние положения элементов колеблющейся среды, т. е. узлы продольных стоячих волн неподвижны, как и при поперечных колебаниях.)

Существенное отличие стоячих волн от бегущих в том, что **в стоячей волне нет переноса энергии, хотя плотность энергии меняется от точки к точке и зависит от времени**. Это очевидно хотя бы потому, что точки узлов волны неподвижны и переноса энергии через них быть не может. Можно рассуждать и иначе: две бегущие волны, образовавшие стоячую, переносят одну и ту же энергию, но в противоположных направлениях, так что оба эти процесса взаимно компенсируются. Подставляя найденное решение (27.53) в выражение (27.39), получаем для плотности энергии стоячей волны:

$$w = \frac{\rho \omega^2 u_{\text{max}}^2}{2} (\cos^2 \omega t \sin^2 kx + \sin^2 \omega t \cos^2 kx). \quad (27.54)$$

Усредняя по времени, находим:

$$\langle w \rangle = \frac{\rho\omega^2 u_{\max}^2}{4} (\sin^2 kx + \cos^2 kx) = \frac{\rho\omega^2 u_{\max}^2}{4} = 2 \frac{\rho\omega^2 A^2}{2}. \quad (27.55)$$

Мы получили, что средняя плотность энергии стоячей волны не зависит от точки наблюдения и равна сумме средних плотностей энергий двух бегущих волн, суперпозицией которых она является,

**Задача 27.5.** Для струны, закрепленной на одном конце, найти точки, в которых плотность энергии колебаний не зависит от времени.

**Решение.** Используя соотношение  $\sin^2 \omega t = 1 - \cos^2 \omega t$ , представим (27.54) в виде

$$w = \frac{\rho\omega^2 u_{\max}^2}{2} [\cos^2 kx + \cos^2 \omega t (\sin^2 kx - \cos^2 kx)]. \quad (27.56)$$

Теперь ясно, что зависимости от времени не будет, если  $\sin^2 kx = \cos^2 kx$ , или  $\operatorname{tg} kx = \pm 1$ , т. е. в точках  $x = (\pi/4 + \pi n)/k$ ,  $n = 0, 1, 2, \dots$  ■

**Задача 27.6.** Найти закон изменения во времени плотности энергии стоячей волны в точках, где смещение струны достигает своих максимальных значений.

**Решение.** Указанные точки имеют координаты  $x_{\max, n} = (\pi/2 + \pi n)/k$ ,  $n = 0, 1, 2, \dots$ . В этих точках  $\sin^2 kx_{\max, n} = 1$ ,  $\cos^2 kx_{\max, n} = 0$ , так что из (27.54) следует:

$$w = \frac{\rho\omega^2 u_{\max}^2}{2} \cos^2 \omega t = 2\langle w \rangle \cos^2 \omega t. \quad (27.57) \quad \blacksquare$$

## Гармоники

На струне, закрепленной на одном конце, могут существовать колебания любых частот. Закрепим теперь другой конец струны в точке с координатой  $x = l$ . Имеем то же решение (27.53), которое должно удовлетворять дополнительному граничному условию  $u(l, t) = 0$ . Это означает, что  $\sin kl = 0$ . На струне длиной  $l$ , закрепленной на концах, могут существовать только стоячие волны с волновыми векторами  $k = k_n = \pi n/l$ , где  $n = 1, 2, 3, \dots$ . Соответствующие длины волн  $\lambda_n = 2\pi/k_n = 2l/n$ . Иными словами, на длине струны должно укладываться целое число полуволен. Стало быть, закрепленная с обоих концов струна может колебаться только с определенными частотами:

$$\nu_n = \frac{v}{\lambda_n} = n \frac{1}{2l} \sqrt{\frac{T}{\rho}}. \quad (27.58)$$

Мы использовали здесь уравнение (27.2) для скорости волн на натянутой струне. Видим, что частота колебаний повышается при: 1) уменьшении длины струны, 2) уменьшении ее толщины (линейной плотности) и 3) увеличении натяжения. Эти закономерности известны каждому, кто держал в руках хотя бы гитару. Колебания с низшей частотой ( $n = 1$ ) называются основной (первой) гармоникой, с последующими частотами — высшими (второй, третьей и т. д.) гармониками.

Аналогичные граничные условия существуют и для колебаний воздуха в трубах духовых музыкальных инструментов. Соответственно в их сигналах так-же присутствуют только вполне определенные частоты. Рассмотрим, например,

органныю трубу длиной  $l$ . Волна давления в ней также может быть описана уравнением вида (27.53):

$$\Delta p(x, t) = \Delta p_{\max} \sin \omega t \sin kx.$$

Если труба открыта с обеих сторон, то давление на концах равно стационарному (атмосферному) и  $\Delta p = 0$  в точках  $x = 0, l$ . Отсюда получаем те же условия: для волновых чисел  $k = k_n = n\pi/l$ , длин волн  $\lambda = \lambda_n = 2l/n$  и частот  $\nu = \nu_n = vn/(2l)$  ( $n = 1, 2, \dots$ ). Если же труба открыта в точке  $x = 0$  и закрыта на другом конце ( $x = l$ ), то на закрытом конце смещение частиц равно нулю, а давление достигает максимума или минимума:  $kl = \pi/2 + \pi n$ . Отсюда следуют несколько иные соотношения:  $k = k_n = (2n + 1)\pi/(2l)$ ,  $\lambda = \lambda_n = 4l/(2n + 1)$  и  $\nu = \nu_n = v(2n + 1)/(4l)$  ( $n = 0, 1, 2, \dots$ ). Первая гармоника для такой трубы возбуждается на частоте  $\nu_1 = v/(4l)$ , что в 2 раза меньше частоты  $v/(2l)$  первой гармоники полностью открытой трубы.

**Задача 27.7.** Нейлоновая гитарная струна имеет линейную плотность  $\rho = 7,2$  г/м и натянута с силой  $T = 150$  Н. Длина струны  $l = 90$  см. Каковы четыре низшие частоты, извлекаемые на такой струне?

**Решение.** Скорость волны на струне равна:

$$v = \sqrt{\frac{T}{\rho}} = \sqrt{\frac{150}{6,08 \cdot 10^{-3}}} = 157,1 \text{ м/с.}$$

Наибольшая длина стоячей волны равна  $\lambda = 2l = 1,8$  м. Отсюда находим самую низкую частоту:  $\nu_1 = v/\lambda = 157,1/1,8 = 87,3$  Гц. (Эта частота соответствует ноте «фа» большой октавы.) Следующие частоты являются целыми кратными  $\nu_1$ :  $\nu_n = n\nu_1$ . Отсюда следует:  $\nu_2 = 2 \times 87,3 = 174,6$  Гц («фа» малой октавы),  $\nu_3 = 3 \times 87,3 = 261,9$  Гц («до» первой октавы) и  $\nu_4 = 4 \times 87,3 = 349,2$  Гц («фа» первой октавы). ■

**Задача 27.8.** Струна звучит на ноте «до» первой октавы. Максимальное отклонение точек струны от положения равновесия равно  $u_{\max} = 2$  мм. Найти максимальную скорость и ускорение точек струны.

**Решение.** Закон колебания струны имеет вид (см. (27.53)):

$$u(x, t) = u_{\max} \sin \omega t \sin kx, \quad (27.59)$$

откуда находим скорость и ускорение точек струны:

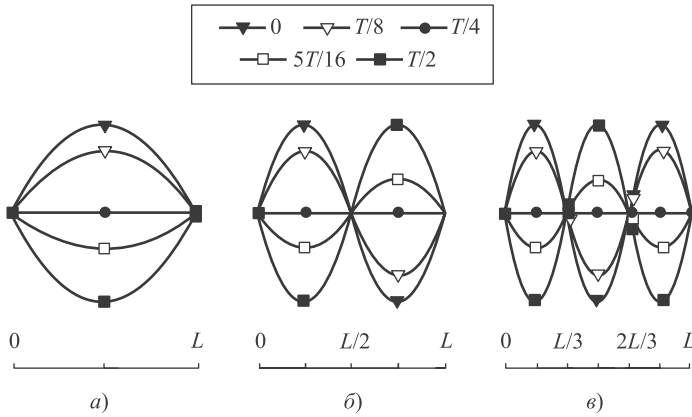
$$\begin{aligned} \dot{u}(x, t) &= \omega u_{\max} \cos \omega t \sin kx, \\ \ddot{u}(x, t) &= -\omega^2 u_{\max} \sin \omega t \sin kx. \end{aligned} \quad (27.60)$$

По условию  $\omega = 2\pi\nu_3 = 1645,6$  Гц (см. задачу 27.7). Максимальные значения скорости и ускорения равны:

$$\begin{aligned} \dot{u}_{\max} &= \omega u_{\max} = 1645,6 \times 2 \cdot 10^{-3} = 3,3 \text{ м/с,} \\ \ddot{u}_{\max} &= \omega^2 u_{\max} = (1645,6)^2 \times 2 \cdot 10^{-3} = 5416 \text{ м/с}^2. \end{aligned} \quad \blacksquare \quad (27.61)$$

## Сложение гармоник

Выше мы уже занимались сложением колебаний, и теперь нам предстоит проделать то же самое, но для каждой точки колеблющейся струны.



**Рис. 27.5.** Колебания струны: первые три гармоники

Движение струны при возбуждении основной (первой) гармоники показано на рис. 27.5, *a*. В начальном положении (линия отмечена черным треугольником) точки струны имеют максимальное отклонение от положения равновесия. Под действием силы упругости они начинают двигаться (на рисунке вертикально вниз). Форма струны через  $1/8$  периода показана линией со светлым треугольником. Через четверть периода струна приходит в положение равновесия (линия с кружком), но ее элементы имеют некоторые скорости, и потому в положении равновесия не задерживаются. Еще через  $1/16$  периода они оказываются в положении, показанном линией со светлым квадратиком, а через половину периода струна снова приобретает максимальное отклонение (линия с черным квадратиком), но в другую сторону, после чего процесс повторяется в обратном направлении. Так происходит колебание в стоячей волне, соответствующее первой гармонике, частоту которой мы обозначаем через  $\nu_1$ .

Но на струне, как мы уже знаем, могут возникать колебания и других гармоник. Вторая гармоника показана на рис. 27.5, *b*. Использование тех же обозначений позволяет не описывать процесс колебаний столь же детально. Заметим, что в этом случае имеется одна неподвижная точка (как мы помним, ее называют узлом) в середине струны. Поскольку узел все равно не движется, можно представить себе, что мы закрепили среднюю точку и тем самым, не изменив частоты издаваемого звука, в 2 раза укоротили струну. Стало быть, частота колебаний ровно в 2 раза превышает частоту основной гармоники:  $\nu_2 = 2\nu_1$ . Этот результат мы уже получали другим путем (см. формулу (27.58) при  $n = 2$ ).

На рисунке 27.5, *в* показаны колебания той же струны при возбуждении третьей гармоники. Здесь уже существует два узла, и издаваемый звук соответствует основному колебанию струны, укороченной в 3 раза (или, что то же самое, второй гармонике струны, длина которой составляет  $2/3$  от исходной):  $\nu_3 = 3\nu_1$ .

На этих рисунках гармоники показаны без учета их амплитуд, т. е. без учета их относительного вклада. На самом деле вклад гармоник может быть различным, так что в общем случае результирующее колебание получается сложением

всех гармоник, каждая из которых представлена уравнением, аналогичным уравнению (27.59):

$$u(x, t) = \sum_{n \geq 1} u_n \cos(2\pi n \nu_1 t) \sin(\pi n x / l). \quad (27.62)$$

Здесь мы уже учли, что частоты и волновые векторы принимают дискретный ряд значений; коэффициенты  $u_n$  — это амплитуды соответствующих гармоник. Заметим, что в множителе с зависимостью от времени мы заменили синус на косинус, т. е. сдвинули точку отсчета времени. Теперь в момент времени  $t = 0$  скорости всех точек струны равны нулю.

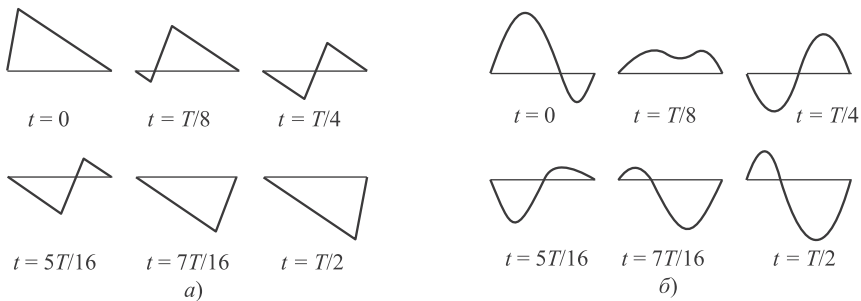
На практике приходится решать обратную задачу: находить коэффициенты  $u_n$  по известному результирующему колебанию. Численные значения  $u_n$  зависят от способа возбуждения струны, например от ее начальной формы  $u(x, 0)$ . Если нам задана функция  $u(x, 0)$ , то для нее из (27.62) при  $t = 0$  следует разложение в ряд Фурье:

$$u(x, 0) = \sum_{n \geq 1} u_n \sin(\pi n x / l). \quad (27.63)$$

Доказана математическая теорема, что коэффициенты  $u_n$  однозначно восстанавливаются по функции  $u(x, 0)$ :

$$u_n = \frac{2}{l} \int_0^l dx u(x, 0) \sin(\pi n x / l). \quad (27.64)$$

От величины примеси высших гармоник зависит форма колеблющейся струны. На рисунке 27.6 показан вид колеблющейся струны в разные моменты времени при двух способах ее возбуждения. В обоих случаях струне придается некая начальная форма, после чего она отпускается «на свободу» (оставаясь закрепленной на концах, разумеется). Мы рассматриваем два случая: а) в начальном положении струна оттягивается за точку, находящуюся на расстоянии  $0,1$  ее длины, считая от ее закрепленного конца; б) начальная форма струны



**Рис. 27.6.** Форма колеблющейся струны в различные моменты времени при разных формах струны в начальном состоянии: а) струна оттянута на расстоянии  $0,1$  ее длины от закрепленного конца; б) струне придана «кусочно-синусоидальная» форма

описывается двумя синусоидами

$$u(x, 0) = \begin{cases} 0, 2 \sin(3\pi x/(2L)), & 0 \leq x \leq 2L/3, \\ -0, 1 \sin(3\pi(x/L - 2/3)), & 2L/3 \leq x \leq L. \end{cases}$$

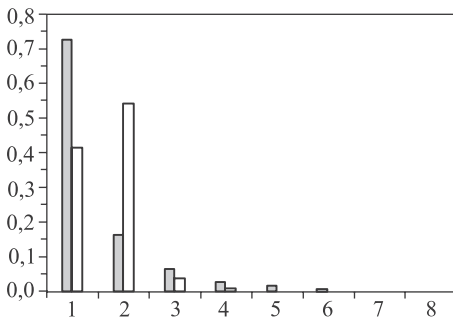
Второй пример носит явно модельный характер и дан для сравнения. Надо сказать, что придание струне какой-то начальной формы — это не единственный способ возбуждения колебаний. Можно, например, задать начальное распределение скоростей (как происходит, скажем, в фортепиано, где молоточек ударяет по струне, находящейся в положении равновесия).

Мы все же ограничимся двумя описанными случаями. На рисунке 27.6 показаны положения точек струны за половину периода колебаний основной гармоникой, после чего процесс повторяется в обратном направлении. За время, равное периоду колебаний  $T_1$  первой гармоникой, струна возвращается в исходное положение. Если струна настроена на ноту «до» первой октавы, то  $T_1 = 3,82$  мс.

## Спектр колебаний

Мы убедились, что любое колебание можно представить в виде суммы (или, как говорят, *суперпозиции*) гармоник. Но причудливое изменение формы струны со временем мало что говорит нам о составе колебаний. Поэтому самой, быть может, важной характеристикой колебания является его *спектр*: набор коэффициентов  $d_n = u_n^2 / \sum_{n=1}^{\infty} u_n^2$ , показывающих относительный вклад каждой гармоникой в полную интенсивность (энергию) колебаний.

Спектр колебания зависит от способа возбуждения струны. Диаграмма на рис. 27.7 демонстрирует результат численного расчета спектра колебаний струны для обоих наших случаев. Темные столбики соответствуют «треугольной» начальной форме струны, а светлыми показан спектр колебаний струны, начальная форма которой составлена из кусков разных синусоид. По оси  $x$  отложены номера гармоник, а высота столбиков по оси  $y$  дает относительную интенсивность соответствующих колебаний (т. е. долю каждой из гармоник в полной интенсивности издаваемого струной звука).



**Рис. 27.7.** Спектр колебаний струны (первые восемь гармоник) при разных способах возбуждения колебаний. Темными столбиками показаны амплитуды обертонов при «треугольной» начальной форме струны, а светлыми — при «кусочно-синусоидальной» форме

Видно, что в первом случае наибольший вклад вносят первые четыре гармоники — на них приходится 97,9% интенсивности звука, вклад следующих четырех гармоник уже невелик (2%), а на все остальные приходится лишь около 0,1% интенсивности. Конкретные числа здесь зависят от способа возбуждения струны, при ином методе они могут измениться. Например, при оттягивании струны за середину все четные гармоники просто отсутствуют в спектре, на долю первой приходится 98,55%, на долю третьей — 1,22%, пятой — 0,16%, а оставшиеся 0,07% — на долю всех остальных.

Во втором случае вклад второй гармоники больше вклада основной, заметны третья и в меньшей степени четвертая гармоники. В сумме вклад первых четырех гармоник составляет около 99,97%, на остальные приходится лишь 0,03%.

Эти примеры демонстрируют общее правило: обычно возбуждаются низшие гармоники колебаний, а влияние высших уменьшается с ростом их номеров. От наличия дополнительных гармоник в спектре колебаний (в музыке их называют *обертонами*) зависит окраска, тембр звука. По-разному воспринимаются одна и та же нота, сыгранная на флейте, скрипке или гобое. Если бы звучала чистая нота, то никакого отличия одного инструмента от другого не было бы. Разнообразие музыкальных звуков мы обязаны обертонам. Степень их присутствия зависит не только от способа возбуждения вибратора, но и (и даже в гораздо большей степени) от резонатора инструмента. Так, при игре на скрипке в образовании звука принимает участие корпус инструмента, колеблющийся под действием колебаний струн, и объем воздуха внутри корпуса. Во-первых, тем самым усиливается звук инструмента: основной источник звука — тонкая струна, и сама по себе она не может привести в движение большую массу воздуха, чтобы звук дошел до слушателя. Во-вторых, на верхней деке скрипки возникают колебания, причем вследствие резонанса некоторые обертоны исходных колебаний струн усиливаются, а другие гасятся. Резонаторы музыкальных инструментов выступают, таким образом, как преобразователи тембра исходного звука. Всем известно, что ценность инструмента, будь то скрипка или рояль, зависит не от качества натянутых в нем струн, а от особенностей строения корпуса, дек. Искусство старых итальянских мастеров Страдивари, Амати, Гварнери, изготовивших прекрасные скрипки, альты и виолончели, состояло, в частности, в том,

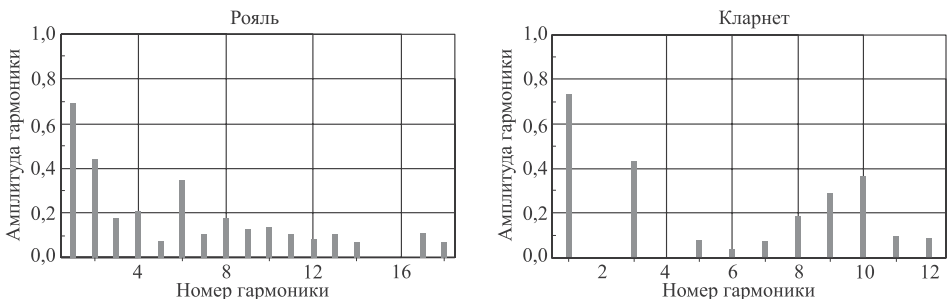


Рис. 27.8. Спектр звука различных музыкальных инструментов



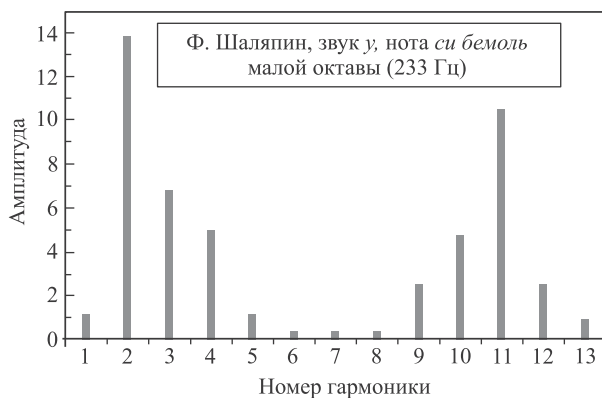


Рис. 27.9. Спектр голоса Ф. И. Шаляпина

что они умели на практике добиваться оптимального спектра колебаний, что мы воспринимаем как божественное звучание их инструментов.

Области усиления высших обертонов называются **формантами**, и они определяют тембр того или иного инструмента. На рисунке 27.8 показано разложение в спектр звуков рояля и кларнета. На спектрах отчетливо видны форманты: скажем, 6-я гармоника для рояля и 8–12-я для кларнета. Именно они создают отличия в звучании этих инструментов.

На рисунке 27.9 показан спектр голоса великого русского певца Ф. И. Шаляпина<sup>1</sup>. Он совершенно необычен. Основной тон у Шаляпина в общем спектре роли не играет (его амплитуда принята на рисунке за единицу). Основная энергия заключена в низкой певческой форманте — 2–4-м обертонах (особенно во второй гармонике с  $\nu_2 \approx 500$  Гц) и в высокой певческой форманте — 10–11-м обертонах (особенно в одиннадцатой гармонике с  $\nu_{11} \approx 2600$  Гц).

## 27.5 Сферические волны

В предыдущих разделах мы рассматривали специальный тип волн: фаза таких волн  $\omega t - kx + \varphi_0$  зависит только от координаты  $x$ . Такие волны называются **плоскими**. Поверхность, на которой фаза колебаний постоянна, называется **фронтом волны**. Для изученных нами волн колебания среды одинаковы во всех точках плоскости, ортогональной направлению распространения волны (мы выбрали его в качестве оси  $x$ ). Иными словами, фронт волны является плоскостью, параллельной плоскости, содержащей оси  $y, z$ . Фронт бегущей плоской волны перемещается с течением времени вдоль оси  $x$  с фазовой скоростью  $v$ .

<sup>1</sup>Л. Б. Дмитриев Основы вокальной методики. М.: Музыка, 1968.

### Трехмерное волновое уравнение

Пусть мы по-прежнему имеем дело с плоской волной, распространяющейся в направлении  $x$ . Такие волны, как мы уже знаем, описываются дифференциальным уравнением (27.15). Повернем координатные оси так, чтобы направление распространения волны задавалось каким-то единичным вектором  $\vec{n}$ . Решение, очевидно, имеет вид

$$u(\vec{r}, t) = A \cos(\omega t - \vec{k} \cdot \vec{r} + \varphi), \quad \vec{k} = k\vec{n}. \quad (27.65)$$

Вектор  $\vec{k}$  называется **волновым вектором**, помимо всего прочего он указывает направление распространения волны. Соотношения между  $\omega$ ,  $k$  и  $v$  остаются прежними. Фронт волны (плоскость, ортогональная вектору  $\vec{k}$ ) движется со скоростью  $v$ , оставаясь параллельным самому себе.

Найдем дифференциальное уравнение, которому удовлетворяет решение (27.65). Дважды дифференцируем по координатам  $x, y, z$ :

$$\begin{aligned} \frac{\partial^2 u(\vec{r}, t)}{\partial x^2} &= -k_x^2 A \cos(\omega t - \vec{k} \cdot \vec{r} + \varphi), \\ \frac{\partial^2 u(\vec{r}, t)}{\partial y^2} &= -k_y^2 A \cos(\omega t - \vec{k} \cdot \vec{r} + \varphi), \\ \frac{\partial^2 u(\vec{r}, t)}{\partial z^2} &= -k_z^2 A \cos(\omega t - \vec{k} \cdot \vec{r} + \varphi). \end{aligned} \quad (27.66)$$

Складывая эти три уравнения, находим:

$$\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) u(\vec{r}, t) = -k^2 A \cos(\omega t - \vec{k} \cdot \vec{r} + \varphi). \quad (27.67)$$

Вторая производная решения по времени дает уравнение:

$$\frac{\partial^2 u(\vec{r}, t)}{\partial t^2} = -\omega^2 A \cos(\omega t - \vec{k} \cdot \vec{r} + \varphi). \quad (27.68)$$

Учитывая соотношение  $v = \omega/k$ , получаем из (27.67), (27.68):

$$\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) u(\vec{r}, t) = \frac{1}{v^2} \frac{\partial^2 u(\vec{r}, t)}{\partial t^2}. \quad (27.69)$$

Выражение в скобках в левой части уравнения является дифференциальным оператором, который называется **лапласианом** и обозначается значком  $\Delta$ . Уравнение в частных производных (27.69) — это волновое уравнение для волн в трехмерном пространстве. Его можно записать в компактной форме:

$$\Delta u(\vec{r}, t) - \frac{1}{v^2} \ddot{u}(\vec{r}, t) = 0. \quad (27.70)$$

(Подчеркнем, что  $\Delta$  не есть греческая буква  $\Delta$  («дельта»), а  $\Delta u$  есть не приращение величины  $u$ , а сумма вторых ее производных по координатам.)

Если волновая функция  $u$  зависит только от одной пространственной координаты (скажем,  $x$ ), то лапласиан превращается во вторую производную по  $x$  и мы возвращаемся к прежней форме волнового уравнения. Но волновое уравнение (27.70) имеет и другие решения, нежели плоские волны. Простым дифференцированием можно убедиться, что *сферическая* волна

$$u(r, t) = \frac{C}{r} \cos(\omega t - kr + \varphi_0) \quad (27.71)$$

удовлетворяет этому уравнению. Фронт волны является сферой с центром в месте расположения источника колебаний ( $\vec{r} = 0$ ), причем радиус сферы увеличивается со скоростью  $v$ , с которой и распространяется фронт сферической волны. Действительно, поверхность постоянной фазы дается уравнением  $\omega t - kr + \varphi_0$ , дифференцируя которое находим  $dr/dt = \omega/k = v$ . Амплитуда сферической волны  $A = C/r$  убывает с увеличением расстояния  $r$  от источника колебаний. Соответственно, интенсивность волны  $I \sim A^2 = C^2/r^2$  убывает по закону обратных квадратов. Это, как и закон Кулона, также связано с трехмерностью нашего пространства. Если среда не поглощает излучение, то поток энергии через поверхность сферы одинаков для сфер любых радиусов, окружающих источник излучения. Поскольку площадь сферы равна  $4\pi r^2$ , энергия, проходящая через единицу площади, обратно пропорциональна  $r^2$ .

## Эффект Доплера для звуковых волн

Стоя у полотна железной дороги, можно наблюдать следующее явление: сигнал приближающегося поезда резко меняет свой тон (частоту) в момент прохождения электрички мимо наблюдателя. Это же явление может заметить наблюдатель, сидящий в поезде и проезжающий мимо сигналищего автомобиля, стоящего на переезде. Изменение наблюдаемой частоты волны (в данном случае звуковой волны) при движении источника или наблюдателя называется *эффектом Доплера* (по имени австрийского физика Х. Доплера, предсказавшего его теоретически в 1842 г.).

### Движущийся наблюдатель, покоящийся источник звука

Пусть имеется источник звука, испускающий сферические звуковые волны. На рисунке 27.10 показано расположение в пространстве четырех последовательных гребней (максимумов) звуковых волн. Пусть волна имеет частоту  $\nu_0$ , тогда расстояние между гребнями равно длине волны  $\lambda_0 = v/\nu_0$ , где  $v$  — скорость звука.

Наблюдатель  $A$  движется прямо на источник звука со скоростью  $v_H$ . Поэтому гребни волн приближаются к нему с увеличенной скоростью  $V = v + v_H$ . С каждым последовательным гребнем волны наблюдатель встретится через время  $T = \lambda_0/V$  после предыдущего. Следовательно, в его восприятии изменяется период колебаний. Наблюдаемая частота волны равна:

$$\nu = \frac{1}{T} = \frac{v + v_H}{\lambda_0},$$

откуда находим:

$$\nu = \nu_0 \frac{v + v_n}{v}. \quad (27.72)$$

Наблюдатель  $B$  удаляется по прямой линии от источника с той же скоростью  $v_n$  (предполагаем, что  $v_n < v$ : наблюдатель, удаляющийся от источника со сверхзвуковой скоростью, вообще не услышит звука). Значит, гребни волн приближаются к нему со скоростью  $V = v - v_n$ , и наблюдаемый период колебаний равен  $T = \lambda_0/V$ . Отсюда получаем для наблюдаемой частоты:

$$\nu = \nu_0 \frac{v - v_n}{v}. \quad (27.73)$$

Наконец, наблюдатель  $P$  движется со скоростью  $v_n$ , составляющей угол  $\theta$  с направлением на источник. На сдвиг частоты влияет только компонента скорости вдоль линии, соединяющей наблюдателя и источник, и мы получаем общую формулу

$$\boxed{\nu = \nu_0 \frac{v + v_n \cos \theta}{v}}. \quad (27.74)$$

Предыдущие формулы (27.72) и (27.73) для частных случаев получаются отсюда при  $\theta = 0$  и  $\theta = \pi$  соответственно.

### Движущийся источник звука, покоящийся наблюдатель

Пусть теперь наблюдатель неподвижен, а звуковые волны испускаются источником, движущимся со скоростью  $v_n$ . На рисунке 27.11 показано расположение в пространстве четырех последовательных гребней звуковой волны, отмеченных прямыми цифрами 1, 2, 3, 4. Эти гребни были испущены, когда источник звука находился в точках, отмеченных курсивными цифрами 1, 2, 3, 4, соответственно. Иначе, точка 1 является центром сферы 1, точка 2 — центром сферы 2 и т. д.

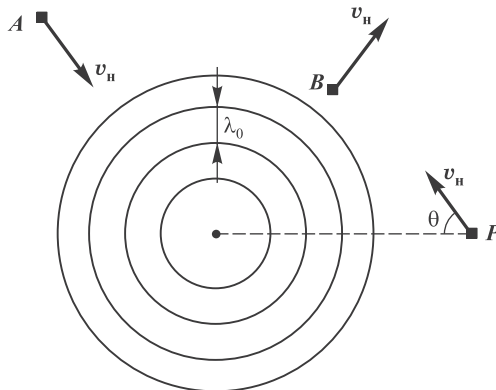


Рис. 27.10. Эффект Доплера при движении наблюдателя

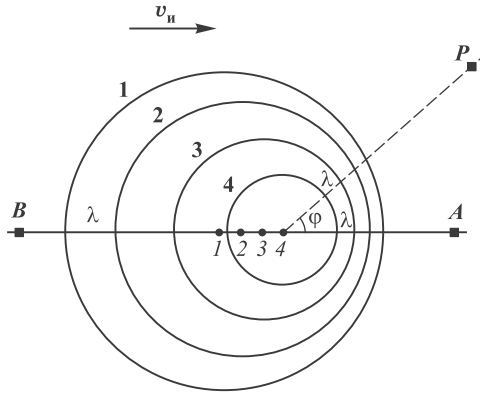


Рис. 27.11. Эффект Доплера при движении источника

Видно, что центры соседних сфер смещаются на расстояние, проходимое источником за период колебаний  $T_0 = 1/\nu_0$ . Это приводит к изменению расстояния между гребнями волн, приходящих к наблюдателю. Следовательно, наблюдатель регистрирует изменившуюся длину волны.

Наблюдатель  $A$  расположен так, что источник движется прямо на него. Для этого наблюдателя расстояние между гребнями волн уменьшается и равно:

$$\lambda = \lambda_0 - v_n T_0 = \lambda_0 - \frac{v_n}{\nu_0}. \quad (27.75)$$

Скорость волны не зависит от движения источника, поскольку определяется свойствами среды. Следовательно, имеем обычную связь между длиной волны и ее фазовой скоростью:  $\lambda = v/\nu$ ,  $\lambda_0 = v/\nu_0$ . Подставляя эти соотношения в (27.75), получаем

$$\frac{v}{\nu} = \frac{v - v_n}{\nu_0},$$

откуда находим частоту  $\nu$  звука, воспринимаемого наблюдателем  $A$ :

$$\nu = \nu_0 \frac{v}{v - v_n}. \quad (27.76)$$

Для наблюдателя  $B$  расстояние между гребнями волн увеличивается и равно:

$$\lambda = \lambda_0 + v_n T_0 = \lambda_0 + \frac{v_n}{\nu_0}.$$

Аналогичные рассуждения приводят к следующему выражению для частоты звуковой волны, воспринимаемой наблюдателем:

$$\nu = \nu_0 \frac{v}{v + v_n}. \quad (27.77)$$

Наконец, для наблюдателя  $P$ , направление на которого составляет угол  $\varphi$  со скоростью источника, общее выражение для частоты имеет вид

$$\nu = \nu_0 \frac{v}{v - v_n \cos \varphi}. \quad (27.78)$$

Предыдущие выражения получаются отсюда при  $\varphi = 0$  и  $\varphi = \pi$ , соответственно.

**Задача 27.9.** Наблюдатель, стоящий на платформе железной дороги, слышит гудок проходящего мимо поезда. Когда поезд приближается, частота звуковых колебаний гудка равна  $\nu_1$ , а когда поезд удаляется, равна  $\nu_2$ . Определить скорость поезда  $V$  и собственную частоту гудка  $\nu_0$ . Скорость звука  $v$  предполагается известной.

**Решение.** При скорости поезда  $V$ , скорости звука  $v$  и собственной частоте колебаний  $\nu_0$  частота  $\nu_1$ , воспринимаемая при приближении поезда, равна:

$$\nu_1 = \nu_0 \frac{v}{v - V}. \quad (27.79)$$

При удалении поезда воспринимаемая частота звука равна:

$$\nu_2 = \nu_0 \frac{v}{v + V}. \quad (27.80)$$

Разделив первое соотношение на второе, получим:

$$\frac{\nu_1}{\nu_2} = \frac{v + V}{v - V}. \quad (27.81)$$

Отсюда находим скорость поезда:

$$V = v \frac{\nu_1 - \nu_2}{\nu_1 + \nu_2}. \quad (27.82)$$

Подставляя скорость поезда в выражение (27.79), получаем

$$\nu_0 = \nu_1 \left(1 - \frac{v}{V}\right) = \nu_1 \left(1 - \frac{\nu_1 - \nu_2}{\nu_1 + \nu_2}\right) = \frac{2\nu_1\nu_2}{\nu_1 + \nu_2}. \quad \blacksquare \quad (27.83)$$

## Движущийся источник звука, движущийся наблюдатель

Из полученных формул можно сделать общие выводы.

1. Для изменения частоты воспринимаемого звука при движении источника или наблюдателя важны не абсолютные значения скоростей, а скорости сближения (удаления), т. е. проекции скоростей на линию, соединяющую источник и наблюдателя.
2. При сближении источника с наблюдателем частота звука увеличивается, при удалении — уменьшается.
3. Если движутся и источник, и наблюдатель, то объединяем формулы (27.72) и (27.76) и получаем общую формулу, описывающую эффект Доплера для звуковых волн:

$$\nu = \nu_0 \frac{v + v_{\text{н}}}{v - v_{\text{и}}}. \quad (27.84)$$

4. В формуле (27.84) в соответствии со сказанным под скоростями  $v_n$  и  $v_n$  надо понимать теперь не абсолютные значения скоростей наблюдателя и источника, а их проекции на линию, соединяющую источник и наблюдателя: положительные знаки скоростей соответствуют сближению, отрицательные — удалению источника и наблюдателя.

Выражение (27.84) явным образом нарушает принцип относительности Галилея. В самом деле, скорость  $v_{\text{отн}}$  сближения источника и наблюдателя есть сумма соответствующих проекций скоростей:  $v_{\text{отн}} = v_n + v_n$ . Согласно принципу относительности, все наблюдаемые эффекты должны зависеть только от  $v_{\text{отн}}$ . Формула же (27.84) позволяет отличить движение наблюдателя от движения источника. Для иллюстрации решим три задачи.

**Задача 27.10.** Сирена полицейской машины, стоящей на обочине дороги, издает сигнал на частоте  $\nu_0 = 1000$  Гц. Звук какой частоты услышит водитель, проезжающий мимо со скоростью  $V = 80$  км/ч?

**Решение.** В данном случае скорость автомобиля  $V = 80$  км/ч = 22,2 м/с — это скорость наблюдателя. При сближении с полицейской машиной водитель воспринимает звук частотой

$$\nu_+ = \nu_0 \frac{v + V}{v} = 1000 \frac{343 + 22,2}{343} \approx 1065 \text{ Гц.}$$

После того как водитель миновал полицейскую машину, воспринимаемая частота становится равной

$$\nu_- = \nu_0 \frac{v - V}{v} = 1000 \frac{343 - 22,2}{343} \approx 935 \text{ Гц.} \quad \blacksquare$$

**Задача 27.11.** Водитель стоящей на обочине дороги машины замечает проезжающий мимо полицейский автомобиль с включенной сиреной. Найти частоту звука, который слышит водитель, если скорость полицейского автомобиля равна  $V = 80$  км/ч. Полицейская сирена — та же самая, что и в предыдущей задаче.

**Решение.** Здесь скорость  $V = 22,2$  м/с — это скорость движения источника. При приближении полиции водитель слышит сигнал частотой

$$\nu_+ = \nu_0 \frac{v}{v - V} = 1000 \frac{343}{343 - 22,2} \approx 1069 \text{ Гц.}$$

При удалении частота воспринимаемого сигнала равна:

$$\nu_- = \nu_0 \frac{v}{v + V} = 1000 \frac{343}{343 + 22,2} \approx 939 \text{ Гц.} \quad \blacksquare$$

**Задача 27.12.** Те же машины движутся навстречу друг другу с равными скоростями  $V = 40$  км/ч = 11,1 м/с. Чему равны частоты звукового сигнала при сближении и при удалении машин?

**Решение.** Применяем формулу (27.84). При сближении воспринимается звук частотой

$$\nu_+ = \nu_0 \frac{v + V}{v - V} = 1000 \frac{343 + 11,1}{343 - 11,1} \approx 1067 \text{ Гц.}$$

При удалении машин сирена для водителя звучит на частоте

$$\nu_- = \nu_0 \frac{v - V}{v + V} = 1000 \frac{343 - 11,1}{343 + 11,1} \approx 937 \text{ Гц.} \quad \blacksquare$$

Во всех трех случаях получились разные результаты, хотя каждый раз относительные скорости сближения (удаления) наблюдателя и источника были теми же самыми. В то же время численные результаты близки друг к другу. Это объясняется тем, что скорости автомобилей в задаче малы по сравнению со скоростью звука. В этом случае в формуле (27.84) можно пренебречь членами  $(v_{н(и)}/v)^2$  и более высоких степеней. Преобразуем (27.84):

$$\begin{aligned}\nu &= \nu_0 \frac{v + v_{н}}{v - v_{и}} = \nu_0 \left(1 + \frac{v_{н}}{v}\right) \cdot \frac{1}{1 - v_{и}/v} = \\ &= \nu_0 \left(1 + \frac{v_{н}}{v}\right) \cdot \left(1 + \frac{v_{и}}{v}\right) \cdot \frac{1}{1 - (v_{и}/v)^2} = \\ &= \nu_0 \left(1 + \frac{v_{н} + v_{и}}{v} + \frac{v_{н}v_{и}}{v^2}\right) \cdot \frac{1}{1 - (v_{и}/v)^2}.\end{aligned}$$

Пренебрегая теперь слагаемыми, содержащими отношения квадратов скоростей, находим приближенное выражение:

$$\nu \approx \nu_0 \left(1 + \frac{v_{н} + v_{и}}{v}\right) = \nu_0 \left(1 + \frac{v_{отн}}{v}\right). \quad (27.85)$$

В (27.85) частота зависит только от относительной скорости источника и наблюдателя. Если бы формула была точна, во всех трех задачах мы получили бы один и тот же ответ для сближения ( $\nu_+$ ) и удаления ( $\nu_-$ ) источника и приемника звука:

$$\begin{aligned}\nu_+ &= 1000 \left(1 + \frac{22,2}{343}\right) = 1065 \text{ Гц}, \\ \nu_- &= 1000 \left(1 - \frac{22,2}{343}\right) = 935 \text{ Гц},\end{aligned}$$

Формула (27.85) удовлетворяет принципу относительности Галилея, но она верна, строго говоря, только при бесконечно большой скорости сигнала. Нарушение принципа относительности Галилея связано с наличием среды. Действительно, при движении тел в среде можно отличить состояние покоя от равномерного прямолинейного движения хотя бы по возникающему при движении ветру (вспомним воображаемый «эфирный ветер» из раздела 9.2 в т. 1). Поэтому системы отсчета при наличии среды неравноправны: из них выделена та, в которой среда как целое покоится.

## Сверхзвуковые скорости

Рассмотрим теперь случай, когда источник звуковых волн движется со скоростью, превышающей скорость звука:  $v_{н} > v$ . Пусть в момент времени  $t = 0$  источник был в точке  $S_0$ , а в момент времени  $t$  он находится в точке  $S_t$  (рис. 27.12). Расстояние между этими точками равно  $v_{н}t$ .



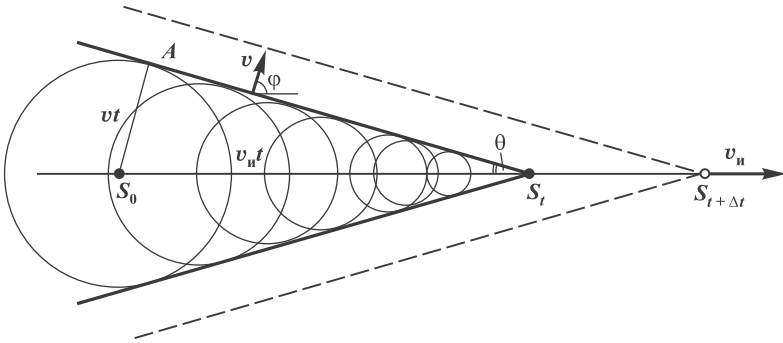


Рис. 27.12. Образование конуса Маха при сверхзвуковом движении источника

В каждой точке своей траектории (для простоты мы рассматриваем равномерное прямолинейное движение) источник испускал сферические звуковые волны. Волна, испущенная в момент  $t = 0$ , к моменту времени  $t$  достигла точки  $A$ . Волны, испущенные на пути от  $S_0$  до  $S_t$ , успели пройти меньшие расстояния. Как видно из рис. 27.12, в каждый момент времени имеется коническая поверхность (ее называют **конусом Маха**), касательная к фронтам всех испущенных сферических волн. Эта коническая поверхность начинается от источника звука, а ее ось совпадает с направлением движения источника. Конус Маха отделяет области пространства, куда дошел звук от источника за время  $t$ , от тех областей, куда звук не успел еще дойти. В некоторый последующий момент времени  $t + \Delta t$  источник переместится в точку  $S_{t+\Delta t}$ . Соответственно переместится и конус Маха, захватив новые области пространства (показано штриховой линией). При движении источника со сверхзвуковой скоростью все возмущения среды, создающие звуковой эффект, находятся только внутри конуса Маха.

Синус угла раствора конуса определяется как отношение расстояния  $vt$ , пройденного звуковой волной за время  $t$ , к расстоянию  $v_n t$ , пройденному источником за то же время:

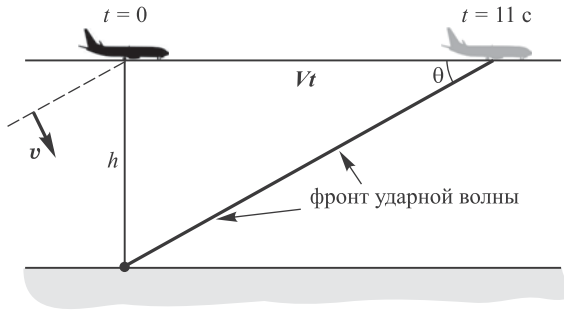
$$\sin \theta = \frac{v}{v_n}. \quad (27.86)$$

Коническую поверхность можно воспринимать как фронт волны (ее называют **ударной**). Направление распространения волны — это нормаль к фронту. Следовательно, ударная волна распространяется под углом  $\varphi = \pi/2 - \theta$  к направлению движения источника. Соответственно формулу (27.86) можно записать в виде

$$\cos \varphi = \frac{v}{v_n}. \quad (27.87)$$

Отношение скорости источника к скорости звука называется **числом Маха**.

**Задача 27.13.** Самолет летит горизонтально на высоте  $h = 5000$  м с постоянной скоростью. Наблюдатель заметил его у себя над головой и засекал время. Звук от самолета появился через  $t = 11$  с после этого. Найти скорость  $V$  самолета. На каком расстоянии по горизонтали находился самолет от наблюдателя в момент, когда последний зарегистрировал приход звука от пролетающего самолета?



**Рис. 27.13.** К задаче 27.13 о пролете сверхзвукового самолета. Пунктирная линия — положение конуса Маха в момент пролета самолета над головой наблюдателя, сплошная — конус Маха в момент, когда звук дошел до наблюдателя

**Решение.** За время  $t$  самолет удалился от наблюдателя на расстояние  $s = Vt$ . Так как в этот момент звук достиг наблюдателя, то точка наблюдения оказалась на конусе Маха (рис. 27.13). Имеем соотношения

$$\sin \theta = \frac{v}{V} \quad \text{и} \quad \operatorname{tg} \theta = \frac{h}{Vt}. \quad (27.88)$$

Отсюда находим:

$$\cos \theta = \frac{\sin \theta}{\operatorname{tg} \theta} = \frac{vt}{h}. \quad (27.89)$$

Используя соотношение  $\sin \theta = \sqrt{1 - \cos^2 \theta}$ , получаем связь скорости самолета с высотой полета и временем  $t$ :

$$\frac{v}{V} = \sqrt{1 - \frac{v^2 t^2}{h^2}},$$

откуда

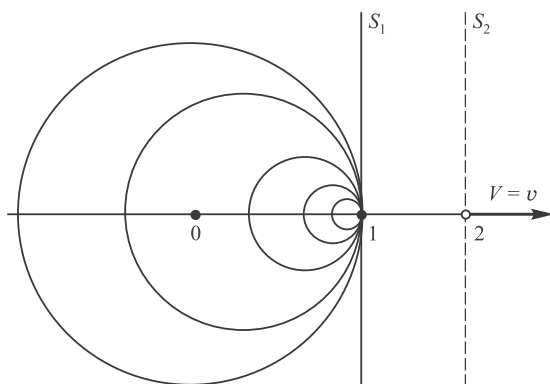
$$V = v \frac{h}{\sqrt{h^2 - v^2 t^2}}. \quad (27.90)$$

Подставляем числовые данные:

$$V = 343 \frac{5000}{\sqrt{5000^2 - (343 \times 11)^2}} = 523 \text{ м/с}. \quad (27.91)$$

Число Маха равно  $V/v = 523/343 = 1,52$ . При такой скорости движения самолет удалился по горизонтали на расстояние  $Vt = 523 \times 11 = 5753 \text{ м} \approx 5,7 \text{ км}$ . Заметим, кстати, что момент прихода к наблюдателю фронта ударной волны воспринимается как резкий хлопок (подобный грому). Бытующее выражение «самолет преодолел звуковой барьер» неверно отражает физический процесс: хлопок, как мы видели, не связан с моментом, когда самолет приобрел сверхзвуковую скорость.

Проанализируем формулу (27.90). Видно, что время задержки  $t$  не может быть больше чем  $t_{\max} = h/v$ . Такая задержка могла бы быть при очень большой скорости самолета ( $V \gg v$ ), когда конус Маха становится крайне узким, почти параллельным направлению движения. Фронту ударной волны, распространяющемуся вертикально вниз, надо при этом пройти расстояние  $h$ , что он и осуществляет за время  $t_{\max}$ .



**Рис. 27.14.** Фронт ударной волны в случае движения источника со скоростью звука. В каждый момент времени источник находится на поверхности всех сферических волн, испущенных ранее. Огибающая этих сферических волн — фронт ударной волны — представляет собой плоскость, ортогональную направлению движения источника

На рисунке 27.14 проиллюстрирован случай, когда самолет летит равномерно и прямолинейно со скоростью звука ( $V = v$ ). Пусть в какой-то момент времени  $t_0$  самолет находился в точке 0, а к моменту времени  $t_1$  переместился в точку 1. Так как самолет летит со скоростью звука, он находится на гребне волны, испущенной им в точке 0. Это же справедливо для всех волн, испущенных им ранее — как до момента  $t_0$ , так и в промежутке времени от  $t_0$  до  $t_1$ . Огибающая всех этих сферических волн — фронт ударной волны — представляет собой плоскость, ортогональную направлению движения источника. В момент времени  $t_1$  она занимает положение  $S_1$ , а к некоторому следующему моменту времени  $t_2$  переместится со скоростью звука в положение  $S_2$ . Из рисунка 27.14 видно, что ударная волна доходит до наблюдателя в тот самый момент, когда самолет оказывается у него над головой. Иными словами, время задержки  $t = 0$ , как это и следует из формулы (27.90) при  $V = v$ . При этом угол раствора конуса Маха становится равным  $\pi/2$  (см. (27.88)).

При дозвуковых скоростях звук опережает летательный аппарат и приходит к наблюдателю раньше самолета. ■

## Контрольные вопросы

1. Запишите уравнение плоской волны.
2. Чему равна скорость распространения волны вдоль натянутой струны?
3. Чему равна скорость распространения звуковой волны в идеальном газе?
4. Чему равна скорость распространения звуковой волны в твердом теле?
5. Как выглядит общее решение волнового уравнения для плоской волны?
6. Запишите решение волнового уравнения для плоской монохроматической волны.
7. Что такое длина волны и частота монохроматической волны? Как они связаны между собой?
8. Может ли дельфин обнаружить муху, упавшую в воду?

9. Волна в первой среде имеет длину  $\lambda_1$  и скорость распространения  $v_1$ . С какой скоростью будет распространяться эта же волна во второй среде, непосредственно примыкающей к первой, если там длина волны становится равной  $\lambda_2$ ?
10. При катании на доске на гребне океанской волны (серфинг) частицы поверхности воды колеблются с периодом  $T = 1,5$  с, а расстояние между гребнями равно  $l = 25$  м. С какой скоростью волна несет человека?
11. От чего зависит среднее значение плотности энергии волны?
12. Откуда следует, что плоская монохроматическая волна переносит энергию?
13. Что такое интенсивность волны?
14. Что такое уровень громкости звуковой волны? В каких единицах он выражается?
15. Что такое стоячая волна? Укажите ее главные отличия от бегущих волн.
16. Найдите зависимость от времени плотности энергии стоячей волны на струне, закрепленной в одной точке. Объясните, почему она отличается от нуля даже в узлах.
17. Что такое гармоники и спектр колебаний?
18. Запишите волновое уравнение для волн, распространяющихся в трехмерном пространстве.
19. Что такое сферические волны? Что представляет собой фронт сферической волны?
20. Как меняется амплитуда сферической волны в зависимости от расстояния до источника?
21. Как меняется плотность энергии сферической волны в зависимости от расстояния до источника?
22. Что такое эффект Доплера? Приведите примеры для звуковых волн.
23. Эффект Доплера для звуковой волны различается в случаях, когда: а) движется источник звука, а наблюдатель покоится; б) когда покоится источник звука и движется наблюдатель. Как объяснить это явное нарушение принципа относительности Галилея?
24. Какие явления возникают при движении источника звука со сверхзвуковой скоростью?



## Часть V

# ОСНОВЫ ВОЛНОВОЙ ОПТИКИ

*Бросая в воду камешки, смотри на круги,  
ими образуемые; иначе такое бросание  
будет пустою забавою.*

---

Козьма Прутков. «Мысли и афоризмы»



## Глава 28

# Волновая теория света и законы геометрической оптики

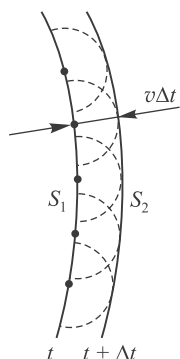
В конце XVII в. на основе многовекового опыта и развития представлений о свете И. Ньютон выдвинул теорию истечения световых частиц (корпускул), которые летят прямолинейно и подчиняются законам механики. По этой теории отражение света уподоблялось отражению упругих шариков при ударе о плоскость, что приводило к известному из школьного курса закону геометрической оптики: «угол падения равен углу отражения». Закон преломления света на границе двух сред объяснялся притяжением световых частиц преломляющей средой, вследствие чего изменялись траектория их движения и скорость. Расчеты приводили к выводу, что скорость световых частиц в более плотных средах больше, чем в воздухе. Однако этот вывод Ньютона в дальнейшем был экспериментально опровергнут. Современник Ньютона Х. Гюйгенс предложил другую теорию света — волновую. Впоследствии волновая теория получила свое обоснование как в экспериментах, так и в рамках теоретических представлений об электромагнитной природе света — в уравнениях Максвелла (см. т. 1, ч. III, гл. 23) и следующих из них свойствах электромагнитных волн (подробнее будут рассмотрены ниже в главе 32). Волновая теория позволила объяснить и изучить такие оптические явления, как дифракция, интерференция и поляризация. Но сначала мы познакомимся в этой главе с тем, как волновая теория света объясняет хорошо известные законы геометрической оптики.

### 28.1 Принцип Гюйгенса

Для анализа распространения света Гюйгенс использовал простой и наглядный метод, названный впоследствии *принципом Гюйгенса*:

♦ *Каждая точка среды, до которой доходит волновое движение, является, в свою очередь, центром вторичных сферических волн. Поверхность, огибающая в некоторый момент времени эти вторичные волны, представляет собой фронт действительно распространяющейся волны в этот момент времени.* ♦





**Рис. 28.1.** Иллюстрация принципа Гюйгенса: каждая точка, до которой дошло волновое движение, может рассматриваться как источник вторичных сферических волн

Иными словами, волновой фронт — это поверхность, до которой в данный момент времени дошел волновой процесс. Отсюда следует, что

♦ *волновой фронт — это поверхность равных фаз, т. е. геометрическое место точек, колеблющихся в одной и той же фазе.* ♦

Пусть в момент времени  $t$  фронт волны занимает положение  $S_1$  (рис. 28.1). Каждую точку этого фронта можно рассматривать как источник вторичных волн, которые в однородной изотропной среде будут представлять собой сферы радиусом  $v\Delta t$ . Таким образом, в момент времени  $t + \Delta t$  фронтом  $S_2$  волны будет огибающая этих вторичных волн.

В сущности, мы уже однажды воспользовались принципом Гюйгенса, когда строили фронт ударной волны — конус Маха — для источника, движущегося со сверхзвуковой скоростью. Точно такое же построение описывает излучение электромагнитных волн (света) заряженной частицей, движущейся быстрее скорости света в среде (эффект Черенкова).

## 28.2 Законы отражения и преломления света

Корпускулярная теория очень просто объясняла явления геометрической оптики, описываемые в терминах распространения световых лучей. С точки зрения волновой теории лучи — это нормали к фронту волны. Принцип Гюйгенса также позволяет объяснить законы геометрической оптики на основе волновых представлений о природе света.

### Закон отражения

Если световые волны достигают границы раздела двух сред, то направление их распространения изменяется, но они остаются в той же среде, т. е. происходит *отражение* света. *Закон отражения света* хорошо известен:

♦ *падающий луч, перпендикуляр к границе раздела двух сред в точке падения и отраженный луч лежат в одной плоскости, причем угол падения равен углу отражения.* ♦

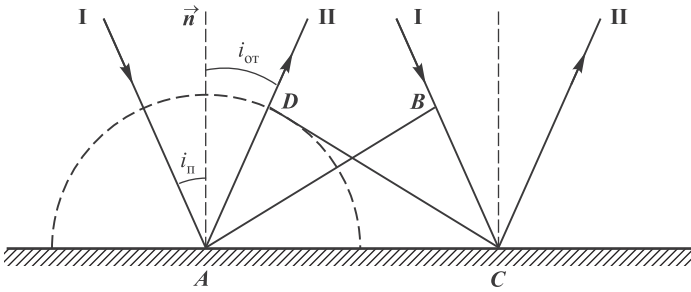


Рис. 28.2. Отражение света на границе двух сред

Допустим, что плоская волна, распространяющаяся в изотропной среде, падает на границу раздела двух сред  $AC$  (рис. 28.2). Достаточно рассмотреть два крайних луча I в падающем пучке. Углом падения называют угол  $i_p$  между нормалью  $\vec{n}$  и падающим лучом I. Плоский фронт  $AB$  падающей волны раньше всего достигнет границы раздела в точке  $A$ , которая станет источником вторичных волн. Согласно принципу Гюйгенса, из нее как из центра будет распространяться в обратном направлении сферическая волна со скоростью  $v$ . Через время  $\Delta t = BC/v$  крайний луч  $BC$  из падающего пучка I придет в точку  $C$ . К этому моменту вторичная сферическая волна, распространяющаяся из точки  $A$ , будет иметь радиус  $v\Delta t = BC$  (т.е.  $AD = BC$ ). Мы знаем теперь положения двух точек фронта отраженной волны —  $C$  и  $D$ . Чтобы не загромождать рисунок, мы не показываем вторичные волны, испущенные точками между  $A$  и  $C$ , но линия  $CD$  является касательной (огibaющей) ко всем этим вторичным волнам. Стало быть,  $CD$  действительно является фронтом отраженной волны. Направление ее распространения (вдоль лучей II) ортогонально фронту  $CD$ . Из равенства треугольников  $ABC$  и  $ADC$  вытекает равенство углов  $DAC$  и  $BCA$ , что, в свою очередь, приводит к закону отражения

$$i_{от} = i_p.$$

Можно сказать иначе: при отражении падающий под углом  $i_p$  луч поворачивается на угол  $\theta = \pi - 2i_p$ .

## Закон преломления

Если световые волны достигают границы раздела двух сред и проникают в другую среду, то направление их распространения также изменяется — происходит *преломление света*. **Закон преломления света (закон Снеллиуса):**

♦ *падающий луч, перпендикуляр к границе раздела двух сред в точке падения и преломленный луч лежат в одной плоскости, причем отношение синуса угла падения луча к синусу угла преломления постоянно для данной пары сред и равно показателю преломления второй среды относительно первой:  $n_{21} = \sin i_p / \sin i_{пр} = n_2/n_1$ .* ♦

Рассмотрим плоскую волну (фронт  $AB$ ), которая распространяется в среде с показателем преломления  $n_1$  вдоль направления I со скоростью  $v_1 = c/n_1$ . Эта

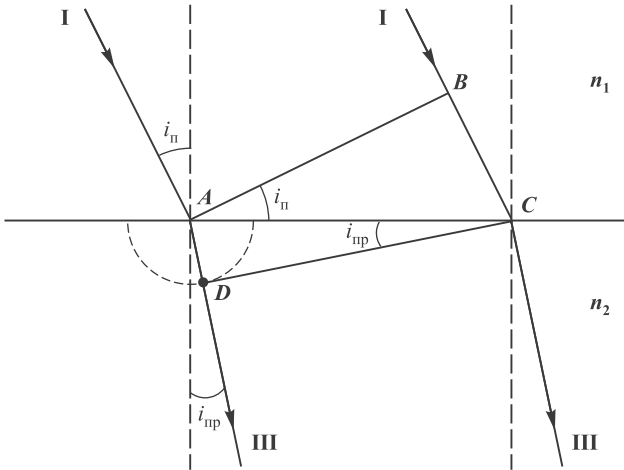


Рис. 28.3. Преломление света на границе раздела двух сред

волна падает на границу раздела со средой, в которой показатель преломления равен  $n_2$ , а скорость распространения луча  $v_2 = c/n_2$  (рис. 28.3).

Время, затрачиваемое падающей волной для прохождения пути  $BC$ , равно  $\Delta t = BC/v_1$ . За это же время фронт волны, возбуждаемой точкой  $A$  во второй среде, достигнет точек полусферы с радиусом  $AD = v_2 \Delta t$ . Согласно принципу Гюйгенса, положение фронта преломленной волны в этот момент времени задается плоскостью  $DC$ , а направление ее распространения — лучом III, перпендикулярным к  $DC$ . Из треугольников  $ABC$  и  $ADC$  следует:

$$AC = \frac{BC}{\sin i_n} = \frac{AD}{\sin i_{np}} \implies \frac{\Delta t \cdot v_1}{\sin i_n} = \frac{\Delta t \cdot v_2}{\sin i_{np}}. \tag{28.1}$$

Таким образом, закон преломления света можно записать так:

$$\boxed{\frac{\sin i_n}{\sin i_{np}} = \frac{v_1}{v_2} = \frac{c}{n_1} \frac{n_2}{c} = \frac{n_2}{n_1} = n_{21}}. \tag{28.2}$$

При преломлении падающий луч поворачивается на угол  $\theta = i_n - i_{np}$ .

Если луч падает из оптически более плотной среды в менее плотную ( $n_1 > n_2$ , или  $n_{21} < 1$ ), мы получаем из (28.2), что выполнение равенства  $\sin i_n = n_{21} \sin i_{np}$  невозможно при  $\sin i_n > n_{21}$ , т. е. при углах падения луча, превышающих значение критического (предельного) угла

$$i_{кр} = \arcsin(n_2/n_1). \tag{28.3}$$

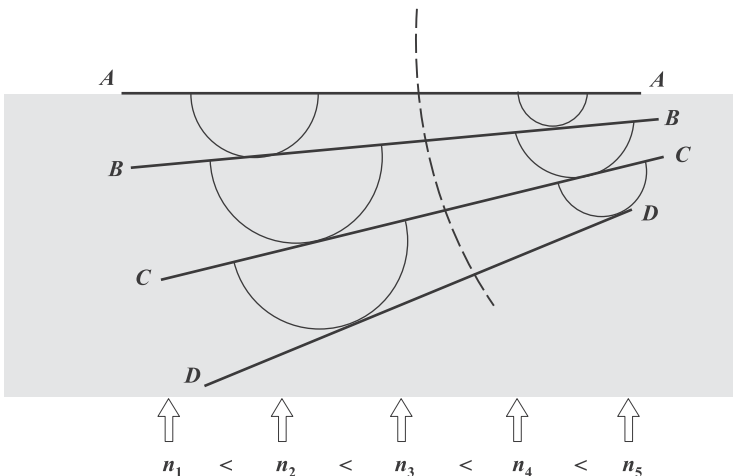
Если угол падения равен **предельному углу** ( $i_n = i_{кр}$ ), то угол преломления  $i_{np} = \pi/2$ , т. е. преломленный луч скользит вдоль поверхности раздела двух сред. Если же угол падения больше предельного угла ( $i_n > i_{кр}$ ), то преломление невозможно — луч целиком отражается в первую среду. Это явление называется

**полным внутренним отражением.** Для границы между водой ( $n_1 = 1,33$ ) и воздухом ( $n_2 = 1$ ) предельный угол  $i_{\text{кр}} \approx 49^\circ$ . Если смотреть из-под воды на ее поверхность, то вместо предметов, находящихся в воздухе, видно зеркальное отражение предметов, находящихся в воде. При полном отражении света от слоев воздуха с разной температурой могут возникнуть миражи. Явление полного внутреннего отражения используется в технике при изготовлении световодов: сердцевина оптоволоконного кабеля сделана из материала с более высоким показателем преломления, чем окружающая его оболочка.

Для еще одной иллюстрации применения принципа Гюйгенса решим качественную задачу.

**Задача 28.14.** На плоскую границу раздела двух сред падает нормально луч света. Показатель преломления среды непрерывно увеличивается от ее левого края к правому (рис. 28.4). Как будет идти луч света в среде?

**Решение.** Пусть фронт волны  $AA$  подошел к границе раздела двух сред. Точки раздела сред можно рассматривать как центры вторичных волн. Проходит какое-то время  $\Delta t_1$ . Испущенные вторичные сферические волны достигают точек на расстоянии  $c\Delta t_1/n$  от фронта  $AA$ . Поскольку показатель преломления среды  $n$  растет слева направо, эти расстояния убывают слева направо. Огибающая к вторичным волнам — новый фронт  $BB$  — повернется. Если теперь взять точки фронта  $BB$  за источники вторичных волн, то за время  $\Delta t_2$  они породят волны, образующие фронт  $CC$ . Он еще более повернут. Его точки порождают фронт  $DD$  и т. д. Проводя нормаль к волновым фронтам в разные моменты времени, получаем путь светового луча в среде с переменным показателем преломления (штриховая линия). Видно, что луч искривляется в сторону увеличения показателя преломления. Аналогия: если притормозить левые колеса автомобиля, он повернется налево. Для света роль «тормоза» выполняет показатель преломления  $n$ . ■



**Рис. 28.4.** Искривление луча света в среде, показатель преломления которой растет слева направо (к задаче 28.14)

Эта задача имеет отношение к явлению, наблюдающемуся на море. Когда ветер дует с берега, иногда возникает так называемая «зона молчания»: звук колокола с судна не достигает берега. Обычно говорят, что звук относится ветром. Но даже при сильном урагане скорость ветра примерно в 10 раз меньше скорости звука, так что «отнести» звук ветер никак не может. Объяснение заключается в том, что скорость встречного ветра у поверхности моря вследствие трения меньше, чем на высоте. Поэтому скорость звука у поверхности больше, и линия распространения звука загибается кверху, не попадая на берег.

## Принцип Ферма

Итак, волновая оптика способна объяснить явления отражения и преломления света столь же успешно, как и геометрическая оптика. В основе последней, трактующей явления на основе законов распространения лучей, лежит *принцип Ферма*:

♦ свет распространяется по такому пути, для прохождения которого требуется минимальное время. ♦

Для прохождения участка пути  $ds$  свету требуется время  $dt = ds/v = n ds/c$ , где  $v = c/n$  — скорость света в среде. Таким образом, время  $t$ , затрачиваемое светом на путь от точки 1 до точки 2, равно:

$$t = \frac{1}{c} \int_1^2 n ds. \quad (28.4)$$

Введем величину с размерностью длины, которая называется *оптической длиной пути*:

$$L = \int_1^2 n ds. \quad (28.5)$$

Пропорциональность  $t$  и  $L$  позволяет сформулировать принцип Ферма следующим образом:

♦ свет распространяется по такому пути, оптическая длина которого минимальна. ♦

Рассмотрим путь света из точки  $S$  в точку  $C$  после отражения на плоскости  $AB$  (рис. 28.5). Непосредственное попадание света из  $S$  в  $C$  невозможно из-за экрана. Нам надо найти точку  $O$ , отразившись в которой луч попадет в точку  $C$ . Среда, в которой проходит луч, однородна. Поэтому минимальность оптического пути сводится к минимальности его геометрического пути. Рассмотрим зеркальное изображение  $S'$  точки  $S$ . Геометрические длины путей  $SOC$  и  $S'OC$  равны. Поэтому минимальность длины  $SOC$  эквивалентна минимальности длины  $S'OC$ . А минимальная геометрическая длина пути из  $S'$  в  $C$  будет соответствовать прямой, соединяющей точки  $S'$  и  $C$ . Пересечение этой прямой с плоскостью раздела двух сред дает положение точки  $O$ . Отсюда следует равенство углов:  $i_{\text{п}} = i_{\text{от}}$ , т. е. закон отражения света.

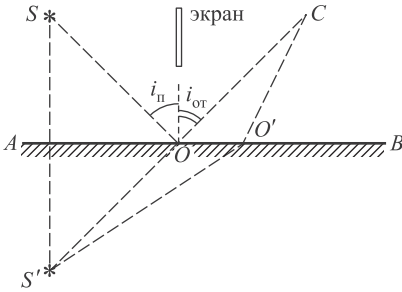


Рис. 28.5. Применение принципа Ферма к отражению света

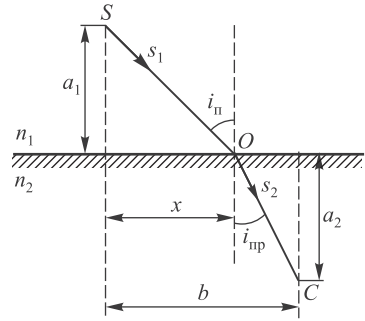


Рис. 28.6. Применение принципа Ферма к преломлению света

Рассмотрим теперь явление преломления света (рис. 28.6). Определим положение точки  $O$ , в которой должен преломиться луч, распространяясь от  $S$  к  $C$ , чтобы оптическая длина пути  $L$  была минимальна. Выражение для  $L$  имеет вид

$$L = n_1 s_1 + n_2 s_2 = n_1 \sqrt{a_1^2 + x^2} + n_2 \sqrt{a_2^2 + (b - x)^2}. \quad (28.6)$$

Найдем величину  $x$ , соответствующую экстремуму оптической длины пути:

$$\frac{dL}{dx} = \frac{n_1 x}{\sqrt{a_1^2 + x^2}} - \frac{n_2 (b - x)}{\sqrt{a_2^2 + (b - x)^2}} = n_1 \frac{x}{s_1} - n_2 \frac{b - x}{s_2} = 0. \quad (28.7)$$

Отсюда следует:  $n_1 \sin i_{\text{н}} - n_2 \sin i_{\text{пр}} = 0$ , или

$$\boxed{\frac{\sin i_{\text{н}}}{\sin i_{\text{пр}}} = \frac{n_2}{n_1}}. \quad (28.8)$$

Мы получили закон преломления света.

Принцип Ферма является частным случаем **принципа наименьшего действия**, имеющего приложения практически ко всем областям физики. Всякий раз из всех возможных движений системы осуществляется то, для которого некая величина (ее называют **действием**) минимальна (точнее, имеет экстремум). В этом проявляется некая «экономность» природы, выбирающей кратчайшие пути для перехода системы из одного состояния в другое.

## Контрольные вопросы

1. Сформулируйте принцип Гюйгенса.
2. Объясните законы отражения и преломления света на основе принципа Гюйгенса.
3. В каких случаях угол преломления равен углу падения?
4. Что такое полное внутреннее отражение света?
5. Чему равен предельный угол при полном внутреннем отражении света?

6. Чем ближе к поверхности Земли, тем больше показатель преломления воздуха. Как это скажется на ходе луча света, падающего на Землю из космоса?
7. Что такое оптическая длина пути?
8. Сформулируйте принцип Ферма.
9. Выведите законы отражения и преломления света на основе принципа Ферма.

## Глава 29

# Интерференция света

*Интерференция волн* — это явление усиления или ослабления колебаний в разных точках пространства, т. е. пространственное перераспределение энергии при наложении двух или нескольких волн. Интерференция возможна для волн любой природы. Интерференция световых волн характеризуется регулярным чередованием в пространстве областей повышенной и пониженной интенсивности света. Как мы убедимся, интенсивность света при интерференции зависит от соотношения между фазами складывающихся колебаний. Для возникновения интерференции необходимо, чтобы накладывающиеся волны были **когерентны** (с этим важным понятием мы познакомимся в данной главе).

### 29.1 Интенсивность света

Для световых волн справедлив принцип суперпозиции — следствие линейности дифференциальных уравнений Максвелла. Так как свет имеет электромагнитную природу, применение принципа суперпозиции означает, что результирующая напряженность электрического (магнитного) поля двух световых волн, приходящих в одну точку, равна векторной сумме напряженностей электрического (магнитного) поля каждой из волн в отдельности.

Мы уже знаем из раздела 27.3, что интенсивность волны пропорциональна квадрату ее амплитуды. Рассмотрим две монохроматические световые волны *одинаковой частоты*, которые накладываются друг на друга и возбуждают в некоторой точке пространства колебания *одинакового направления* (на самом деле физическая природа волн нам сейчас не важна):

$$\begin{aligned} E_1 &= A_1 \cos(\omega t + \alpha_1), \\ E_2 &= A_2 \cos(\omega t + \alpha_2). \end{aligned} \tag{29.1}$$

Фазы  $\alpha_i$  колебаний в данной точке зависят, вообще говоря, от пути, пройденного световой волной от источника до точки наблюдения, а они могут быть разными для обеих волн. Амплитуда  $A$  результирующего колебания в данной точке определяется с помощью векторной диаграммы (см. разд. 24.4). Поскольку интенсивности световых волн пропорциональны квадратам их амплитуд ( $I \propto A^2$ ,



$I_i \propto A_i^2$ ), применение формулы (24.47) дает для интенсивности результирующей волны выражение

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(\alpha_1 - \alpha_2). \quad (29.2)$$

◆ Если разность фаз возбуждаемых волнами колебаний остается постоянной во времени, то волны называются когерентными<sup>1</sup>. ◆

Для когерентных волн косинус разности фаз в формуле (29.2) для интенсивности результирующего колебания имеет постоянное во времени значение (но свое для каждой точки пространства). Последнее слагаемое в выражении (29.2) для результирующей интенсивности носит название интерференционного члена. Из-за наличия этого осциллирующего в пространстве интерференционного члена, который при наложении когерентных волн не изменяется во времени, происходит *устойчивое во времени* перераспределение светового потока в пространстве, в результате чего в одних местах возникают максимумы, а в других — минимумы интенсивности света. Если интенсивности обеих интерферирующих волн одинаковы ( $I_1 = I_2$ ), то в максимумах  $I = 4I_1$ , а в минимумах  $I = 0$ .

При сложении некогерентных волн в данную точку пространства приходят волны, разность фаз которых зависит от времени. Если эта разность фаз меняется случайным образом, то усредненное по времени значение косинуса в интерференционном члене соотношения (29.2) равно нулю. Тогда наблюдаемые средние значения интенсивностей просто складываются и во всех точках пространства

$$I = I_1 + I_2. \quad (29.3)$$

При равенстве интенсивностей приходящих волн получаем  $I = 2I_1$ . Когда мы включаем две одинаковые лампочки и, грубо говоря, помещение освещается в 2 раза ярче, чем одной из них, то это — проявление соотношения (29.3). Таким образом, *необходимым условием наблюдения интерференции волн является их когерентность*.

## 29.2 Когерентность световых волн

Монохроматическая световая волна описывается выражением

$$E = A \cos(\omega t - \vec{k} \cdot \vec{r} + \alpha), \quad (29.4)$$

где  $A$ ,  $\omega$ ,  $\alpha$  и  $k = \omega/v$  являются постоянными величинами. Однако всякая реальная световая волна образуется наложением колебаний различных частот, заключенных в конечном интервале  $\Delta\omega$ . Согласно формуле  $k = \omega/v$  разбросу частот  $\Delta\omega$  соответствует разброс значений волнового вектора  $k$ . Следует отметить, что разброс волнового вектора связан также с разбросом направлений распространения волн, который характеризуется величиной  $\Delta\vec{k}$ . Сначала обсудим *временную когерентность*, которая связана с разбросом частот  $\Delta\omega$ .

<sup>1</sup>Когерентность — согласованное протекание нескольких колебательных или волновых процессов.

Рассмотрим случай наложения в некоторой точке пространства двух световых колебаний с несколько различающимися частотами  $\omega_1$  и  $\omega_2$ :

$$\begin{aligned} E_1 &= A_1 \cos(\omega_1 t + \alpha_1) \\ E_2 &= A_2 \cos(\omega_2 t + \alpha_2). \end{aligned}$$

Интерференционный член

$$\cos(\Delta\omega t + \Delta\alpha)$$

при сделанных предположениях будет зависеть от времени и разности частот  $\Delta\omega = |\omega_1 - \omega_2|$ .

Всякий оптический прибор, с помощью которого наблюдается интерференция света (фотопленка, человеческий глаз и т. д.), обладает определенной инертностью, которая характеризует время  $t_{\text{приб}}$  регистрации прибором интерференционной картины. При этом оптический прибор регистрирует картину, усредненную по промежутку времени  $t_{\text{приб}}$ . Если за это время интерференционный множитель  $\cos(\Delta\omega t + \Delta\alpha)$  с равной вероятностью принимает любые значения в интервале от  $-1$  до  $+1$ , то среднее значение интерференционного слагаемого будет равно нулю. Интерференционная картина не будет видна, т. е. регистрируемая прибором интенсивность окажется равной сумме интенсивностей, создаваемых в данной точке каждой волной в отдельности. Если же за время  $t_{\text{приб}}$  значение косинуса остается практически неизменным, то прибор зарегистрирует интерференцию. Таким образом, для характеристики степени когерентности свойств световых волн вводится **время когерентности**  $t_{\text{ког}}$ , определяемое как время, за которое случайное изменение фазы волны не превышает величину  $\sim\pi$ :  $t_{\text{ког}} \sim \pi/\Delta\omega$ . При  $t_{\text{приб}} \gg t_{\text{ког}}$  прибор не зафиксирует интерференцию, а при  $t_{\text{приб}} \ll t_{\text{ког}}$  прибор обнаружит интерференционную картину. За время когерентности  $t_{\text{ког}}$  волна распространяется на расстояние  $l_{\text{ког}} = ct_{\text{ког}}$ , называемое **длиной когерентности**.

Для наблюдения интерференционной картины обычно используют пучки света от одного источника, но прошедшие разные расстояния до точки наблюдения. Это означает, что интерферируют волны, испущенные источником в разное время. Если частота источника «плавает» или по каким-либо иным причинам происходит случайное изменение во времени фазы волны, то при разности хода волн до точки наблюдения  $\Delta l \geq l_{\text{ког}}$  разница во времени  $\Delta t$  испускания волн будет больше времени когерентности  $\Delta t_{\text{ког}}$ , что означает невозможность наблюдения интерференции.

В качестве примера укажем типичные значения длины когерентности для естественного оптического источника с узкополосным светофильтром с шириной полосы  $\Delta\lambda \approx 1$  нм в центре видимого диапазона ( $\lambda = 500$  нм) и для газового лазера — источника оптического излучения с высокой временной когерентностью, ширина полосы которого на два-три порядка меньше. В первом случае оценки для длины когерентности дают значение  $l_{\text{ког}} \approx 250$  мкм, а во втором случае (для лазера) —  $l_{\text{ког}} \approx 25$  см. Таким образом, наблюдение интерференционной картины от обычных оптических источников возможно лишь при малых разностях хода волн, например при интерференции в тонких пленках, в то время как использование лазерного излучения существенно упрощает постановку экспериментов.

В идеализированном случае, когда складываются монохроматические волны со строго фиксированными частотами ( $\Delta\omega = 0$ ), время и длина когерентности становятся бесконечно большими. Естественно, в таких условиях интерференционная картина наблюдалась бы при любых разностях хода.

Изменение разности фаз колебаний может происходить не только из-за разброса частот  $\Delta\omega$ , но и вследствие разброса волновых векторов  $\Delta\vec{k}$ . Поэтому наряду с временной когерентностью, определяемой временем когерентности, вводится понятие *пространственной когерентности*. Возникновение в некоторой точке пространства колебаний, возбуждаемых волнами с целым набором различных значений  $\vec{k}$ , имеет место, если эти волны испускаются различными участками протяженного источника света.

Рассмотрим для определенности свечащийся диск  $AB$ , который из точки  $M$  виден под углом  $\Delta\varphi$  (рис. 29.1). Угол  $\Delta\varphi$  характеризует разброс волновых векторов  $\Delta\vec{k}$ . Таким образом, в фазу электромагнитной волны

$$\alpha(t) = \omega t - \vec{k} \cdot \vec{r} + \alpha_0$$

надо подставить выражения

$$\begin{aligned} \vec{k} &= \vec{k}_0 + \Delta\vec{k}, & k_0 &= \frac{\omega}{v} = \frac{2\pi}{\lambda}, \\ |\Delta\vec{k}| &= k_0 \Delta\varphi. \end{aligned} \quad (29.5)$$

Тогда

$$\alpha(t) = \omega t - \vec{k}_0 \cdot \vec{r} - k_0 \rho \Delta\varphi + \alpha_0, \quad (29.6)$$

где  $\rho$  — проекция радиуса-вектора на направление вектора  $\Delta\vec{k}$ . Последний, как видно из рис. 29.1, можно считать параллельным протяженному источнику, т. е. фронту волны.

Следовательно, фаза колебаний при переходе от одной точки волновой поверхности к другой изменяется беспорядочным образом. Введем расстояние  $\rho_{\text{ког}}$ , при смещении на которое вдоль волновой поверхности случайное изменение фазы достигает значения  $\sim\pi$ :  $k_0 \Delta\varphi \rho_{\text{ког}} \sim \pi$ , откуда

$$\rho_{\text{ког}} \sim \frac{\pi}{k_0 \Delta\varphi} \sim \frac{\lambda}{2\Delta\varphi}. \quad (29.7)$$

Расстояние  $\rho_{\text{ког}}$  называется *длиной пространственной когерентности* или *радиусом когерентности*.

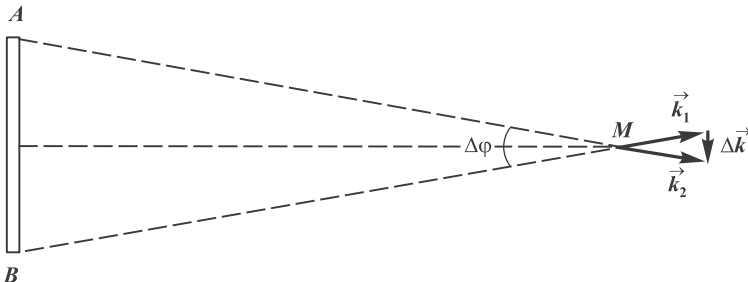


Рис. 29.1. Пространственная когерентность протяженного источника света: угол  $\Delta\varphi$  характеризует разброс волновых векторов  $\Delta\vec{k}$

Пусть длина  $l_{\text{приб}}$  характеризует пространственное разрешение прибора (например, фотопленки или человеческого глаза). Интерференционный член усредняется тогда по части пространства с линейными размерами порядка  $l_{\text{приб}}$ . При  $l_{\text{приб}} \gg r_{\text{ког}}$  среднее значение косинуса равно нулю, что не позволяет наблюдать интерференцию. В противоположном случае, когда  $l_{\text{приб}} \ll r_{\text{ког}}$ , наблюдается четкая интерференционная картина.

Приведем пример. Угловой размер Солнца  $\Delta\varphi \sim 0,01$  рад (порядка полуградуса), а основная длина волны солнечного света  $\lambda \sim 0,5$  мкм. Следовательно, радиус когерентности приходящих от Солнца световых волн по порядку величины равен

$$\rho_{\text{ког}} \sim \frac{\lambda}{2\Delta\varphi} = 25 \text{ мкм} = 0,025 \text{ мм.}$$

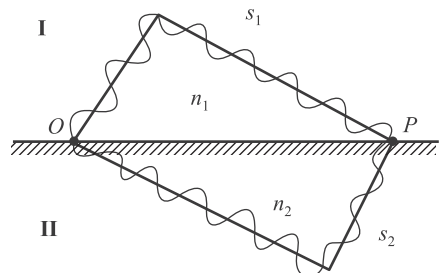
При таком малом радиусе когерентности невозможно непосредственно наблюдать интерференцию солнечных лучей, поскольку разрешающая способность человеческого глаза составляет лишь 0,1 мм. Однако в 1807 г. Т. Юнг провел первое наблюдение интерференции именно с солнечным светом. Для этого он пропускал солнечные лучи в темную комнату через маленькое отверстие, сделанное тонкой иглой. Отверстие на несколько порядков уменьшало угловой размер  $\Delta\varphi$  источника света и соответственно увеличивало радиус когерентности.

### 29.3 Интерференция света от двух источников

Обычно для получения когерентных световых волн тем или иным способом разделяют на две части волну, излучаемую одним источником. После прохождения различных оптических путей эти две части одной волны накладываются друг на друга (рис. 29.2).

Предположим, что разделение на две когерентные волны происходит в некоторой точке  $O$ , лежащей на границе раздела двух сред I и II. До точки  $P$ , в которой наблюдается интерференционная картина, одна волна проходит путь  $s_1$  в среде с показателем преломления  $n_1$ , а другая волна — путь  $s_2$  в среде с показателем преломления  $n_2$ . Если начальные фазы обоих колебаний равны нулю, то в точке  $P$  волны возбуждают колебания:

$$A_{01} \cos \omega \left( t - \frac{s_1}{v_1} \right) \quad \text{и} \quad A_{02} \cos \omega \left( t - \frac{s_2}{v_2} \right), \quad (29.8)$$



**Рис. 29.2.** Оптическая разность хода волн  $\Delta L$  в точке  $P$  возникает и в том случае, если две волны проходят одинаковые пути  $s_1 = s_2$  в средах с разными показателями преломления:  $n_1 \neq n_2$

где  $v_1 = c/n_1$ ,  $v_2 = c/n_2$  — фазовые скорости первой и второй волн соответственно. Разность фаз колебаний в точке  $P$  равна:

$$\alpha_2 - \alpha_1 = \omega \frac{s_1}{v_1} - \omega \frac{s_2}{v_2} = \frac{\omega}{c} (n_1 s_1 - n_2 s_2) = \frac{\omega}{c} (L_1 - L_2). \quad (29.9)$$

Выражая циклическую частоту через длину волны  $\lambda$  в вакууме, находим:

$$\alpha_2 - \alpha_1 = 2\pi \frac{\Delta L}{\lambda}. \quad (29.10)$$

♦ *Разность  $\Delta L = L_1 - L_2$  оптических длин проходимых волнами путей называется оптической разностью хода.* ♦

Запишем интенсивность результирующей волны в точке  $P$  в виде (29.2):

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(\alpha_1 - \alpha_2). \quad (29.11)$$

Если оптическая разность хода равна целому числу длин волн в вакууме (или, что то же самое, четному числу полуволен), т. е.

$$\Delta L = \pm m\lambda = \pm 2m \frac{\lambda}{2}, \quad m = 0, 1, 2, \dots, \quad (29.12)$$

то колебания в точке  $P$  находятся в одинаковой фазе,  $\cos(\alpha_1 - \alpha_2) = 1$  и наблюдается **интерференционный максимум**.

Если оптическая разность хода равна нечетному числу полуволен, т. е.

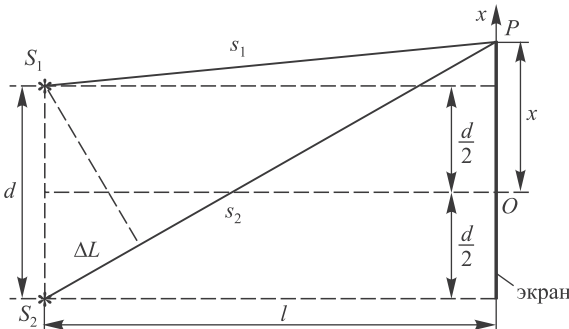
$$\Delta L = \pm(2m + 1) \frac{\lambda}{2}, \quad m = 0, 1, 2, \dots, \quad (29.13)$$

то колебания, возбуждаемые в точке  $P$  обеими волнами, находятся в противофазе и  $\cos(\alpha_1 - \alpha_2) = -1$ , т. е. наблюдается **интерференционный минимум**.

Установив общие закономерности, обратимся к конкретному примеру интерференции, когда источниками света служат две узкие параллельные щели, расположенные достаточно близко друг к другу. Пусть эти щели  $S_1$  и  $S_2$  разделены расстоянием  $d$  (рис. 29.3), а интерференция наблюдается в некоторой точке  $P$  экрана, расположенного на расстоянии  $l$  от источников света ( $l \gg d, \lambda$ ). Начало отсчета оси  $Ox$  выбрано в точке  $O$ , симметричной относительно щелей.

Интенсивность в точке  $P$ , расположенной на расстоянии  $x$  от начала отсчета, определяется оптической разностью хода ( $n = 1$ ):

$$\Delta L = s_2 - s_1. \quad (29.14)$$



**Рис. 29.3.** Интерференция от двух когерентных источников света

Видно, что

$$s_1^2 = l^2 + \left(x - \frac{d}{2}\right)^2, \quad s_2^2 = l^2 + \left(x + \frac{d}{2}\right)^2, \quad (29.15)$$

откуда

$$s_2^2 - s_1^2 = 2xd, \quad \text{или} \quad (s_2 - s_1)(s_1 + s_2) = 2xd, \quad (29.16)$$

т. е.

$$\Delta L = \frac{2xd}{s_1 + s_2}. \quad (29.17)$$

Поскольку  $l \gg d$ , запишем:  $s_1 + s_2 \approx 2l$ . Следовательно,

$$\Delta L = x \frac{d}{l}. \quad (29.18)$$

Учитывая (29.12), (29.13), приходим к выводу, что в точках с координатами

$$x_{\max} = \pm m \lambda \frac{l}{d}, \quad m = 0, 1, 2, \dots \quad (29.19)$$

будут наблюдаться максимумы, а в точках

$$x_{\min} = \pm \left(m + \frac{1}{2}\right) \lambda \frac{l}{d}, \quad m = 0, 1, 2, \dots \quad (29.20)$$

— интерференционные минимумы.

Расстояние между двумя соседними максимумами (или минимумами), называемое *шириной интерференционной полосы*, равно:

$$\Delta x = \lambda \frac{l}{d}, \quad (29.21)$$

причем  $\Delta x$  не зависит от порядка интерференции (величины  $m$ ) и (при сделанных выше предположениях) является постоянной величиной при заданных  $l, d, \lambda$ .

## 29.4 Способы наблюдения интерференции света

Для получения интерференционной картины необходимы, как мы уже знаем, когерентные световые пучки, для создания которых применяются различные искусственные приемы. До появления лазеров во всех устройствах для наблюдения интерференции света когерентные пучки получали разделением и последующим сведением световых лучей, исходящих из одного и того же источника. Практически это можно осуществить с помощью экранов со щелями, зеркал и преломляющих тел (призм). Обсудим некоторые из таких способов.

**Метод Юнга.** Источником света служит ярко освещенная щель  $S$ , от которой свет падает на две равноудаленные щели  $s_1$  и  $s_2$ , параллельные щели  $S$  (рис. 29.4). Таким образом, щели  $s_1$  и  $s_2$  являются источниками когерентных пучков света. Интерференционная картина может наблюдаться на экране, расположенном на некотором расстоянии от щелей.

**Бизеркало Френеля.** Классическим устройством для наблюдения интерференции света является бизеркало Френеля (рис. 29.5). Свет, излучаемый источником  $S$ , отражается от двух зеркал, расположенных под углом, близким к  $180^\circ$  (угол  $\varphi$  достаточно мал). В результате получаются два световых пучка,

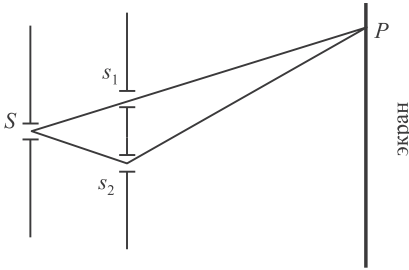


Рис. 29.4. Метод Юнга наблюдения интерференции

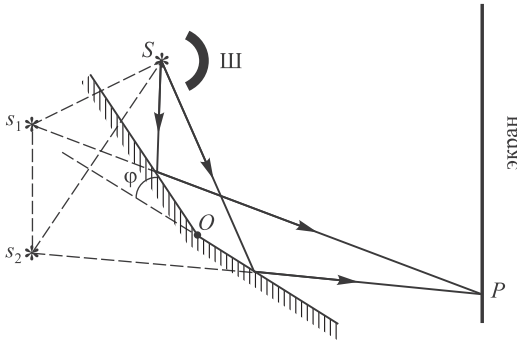


Рис. 29.5. Бизеркало Френеля

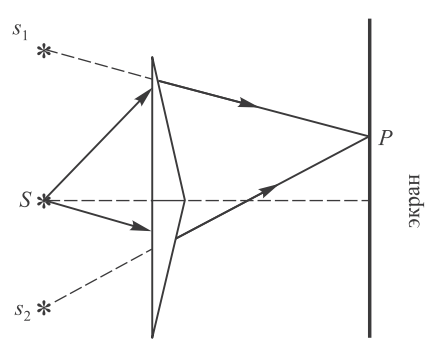


Рис. 29.6. Бипризма Френеля

которые распространяются от двух мнимых источников  $s_1$  и  $s_2$ , дающих когерентные пучки света, так как эти источники являются изображениями одного и того же действительного источника  $S$ . При этом лучи, идущие от  $s_1$  и  $s_2$  к экрану, пройдя различные пути, складываются в интерференционную картину. (Защитная ширма  $\text{Ш}$  преграждает свету непосредственный путь от источника  $S$  к экрану.)

**Бипризма Френеля.** Бипризма Френеля (рис. 29.6) представляет собой две одинаковые призмы с малым преломляющим углом, сложенные основаниями так, что образуется общая плоская грань. Свет от источника  $S$  преломляется в обеих призмах, в результате чего за бипризмой распространяются когерентные световые волны, исходящие из двух мнимых источников  $s_1$  и  $s_2$ . Таким образом, на экране происходит наложение когерентных световых пучков и наблюдается интерференционная картина.

## 29.5 Интерференция света в тонких пленках

Радужная окраска мыльных пузырей или бензиновых пленок на воде возникает в результате интерференции солнечного света, отраженного двумя поверхностями пленки.

Пусть на плоскопараллельную прозрачную тонкую пленку с показателем преломления  $n$  и толщиной  $d$  под углом  $i_n$  падает плоская монохроматическая волна

длиной  $\lambda$  (рис. 29.7). Интерференционная картина в отраженном свете возникает на поверхности пленки из-за сложения двух волн, отраженных от верхней и нижней поверхностей. Рассмотрим сложение волн, выходящих из точки  $C$ . Плоскую волну можно представить как пучок параллельных лучей. Один из лучей пучка (2) непосредственно попадает в точку  $C$  и отражается ( $2'$ ) в ней вверх под углом, равным углу падения  $i_{\text{п}}$ . Другой луч (1) попадает в точку  $C$  более сложным путем: сначала он преломляется в точке  $A$  и распространяется в пленке, затем отражается от нижней поверхности в точке  $O$  и, наконец, выходит, преломившись, наружу ( $1'$ ) в точке  $C$  под углом, равным углу падения  $i_{\text{п}}$ . Таким образом, в точке  $C$  пленка отбрасывает вверх два параллельных пучка света, из которых один образовался за счет отражения от нижней поверхности пленки, другой — вследствие отражения от верхней поверхности пленки. (Пучки, возникающие в результате многократного отражения от поверхностей пленки, не рассматриваются ввиду их малой интенсивности.)

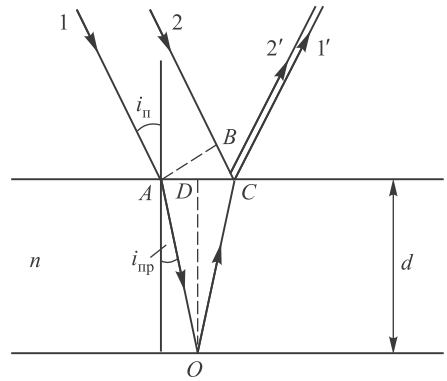


Рис. 29.7. Интерференция света в тонкой пленке

Показатель преломления воздуха полагаем равным единице. Оптическая разность хода, приобретаемая пучками 1 и 2 до того, как они сойдутся в точке  $C$ , равна:

$$\Delta L = (AO + OC)n - BC. \tag{29.22}$$

Учитывая, что  $AO = OC = OD / \cos i_{\text{пр}} = d / \cos i_{\text{пр}}$  и  $\angle BAC = i_{\text{п}}$ , а следовательно,  $BC = AC \cdot \sin i_{\text{п}}$ ,  $AC = 2d \operatorname{tg} i_{\text{пр}}$ , находим:

$$\Delta L = 2n \frac{d}{\cos i_{\text{пр}}} - 2d \operatorname{tg} i_{\text{пр}} \cdot \sin i_{\text{п}} = 2d \frac{n^2 - n \sin i_{\text{п}} \sin i_{\text{пр}}}{n \cos i_{\text{пр}}}. \tag{29.23}$$

Используем закон преломления света:

$$n = \frac{\sin i_{\text{п}}}{\sin i_{\text{пр}}} \quad \text{и} \quad n \cos i_{\text{пр}} = \sqrt{n^2 - n^2 \sin^2 i_{\text{пр}}} = \sqrt{n^2 - \sin^2 i_{\text{п}}}. \tag{29.24}$$

Таким образом,

$$\Delta L = 2d \sqrt{n^2 - \sin^2 i_{\text{п}}}. \tag{29.25}$$

В оптической разности хода  $\Delta L$  следует учесть потерю полуволны, что связано со скачком фазы на  $\pi$  для волны, отраженной от оптически более плотной среды в точке  $C$ . Аналогичный скачок фазы происходит при отражении колебаний струны от закрепленного конца. В точке  $O$  отражение происходит от оптически менее плотной среды, так что скачка фазы не происходит. В итоге между лучами 1 и 2 возникает дополнительная разность фаз  $\pi$ , которую можно учесть, если из  $\Delta L$  вычесть половину длины волны в вакууме.



Следовательно, при выполнении соотношения

$$\Delta L = 2d\sqrt{n^2 - \sin^2 i_{\text{п}}} - \frac{\lambda}{2} = m\lambda, \quad m = 0, 1, 2, \dots \quad (29.26)$$

получается **максимум** интерференции в отраженном свете, а в случае

$$\Delta L = 2d\sqrt{n^2 - \sin^2 i_{\text{п}}} - \frac{\lambda}{2} = \left(m + \frac{1}{2}\right)\lambda, \quad m = 0, 1, 2, \dots \quad (29.27)$$

в отраженном свете наблюдается **минимум**. Таким образом, при падении солнечного света на бензиновую пленку на воде в зависимости от угла зрения и толщины пленки наблюдается радужная окраска пленки, свидетельствующая об усилении световых волн с определенными значениями  $\lambda$ . Это так называемые **полосы равного наклона**. Следует отметить, что интерференция от тонких пленок может наблюдаться не только в отраженном, но и в проходящем свете.

**Задача 29.15.** На мыльную пленку ( $n = 1,3$ ), находящуюся в воздухе, падает по нормали пучок белого света. При какой наименьшей толщине  $d$  пленки отраженный свет с длиной волны  $\lambda = 0,55$  мкм окажется максимально усиленным в результате интерференции?

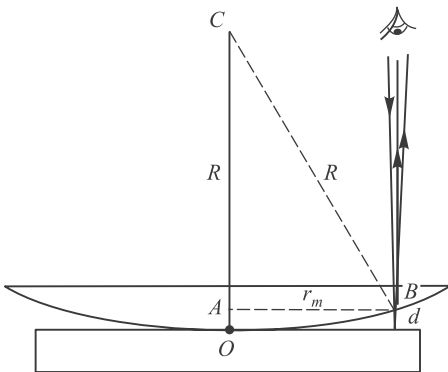
**Решение.** Из условия интерференционного максимума (см. формулу (29.26)) находим для толщины пленки выражение

$$d = \lambda \frac{2m + 1}{4n}$$

(угол падения  $i_{\text{п}} = 0$ ). Минимальное значение  $d$  получается при  $m = 0$ :

$$d_{\text{min}} = \frac{\lambda}{4n} = \frac{0,55}{4 \times 1,3} = 0,1 \text{ мкм. } \blacksquare$$

## 29.6 Полосы равной толщины. Кольца Ньютона



**Рис. 29.8.** Наблюдение колец Ньютона

Рассмотрим отражение света от соприкасающихся друг с другом плоскопараллельной толстой стеклянной пластинки и плоско-выпуклой линзы с большим радиусом кривизны (рис. 29.8). Роль тонкой пленки, от поверхностей которой отражаются когерентные волны, выполняет воздушный зазор между пластинкой и линзой. (Вследствие большой толщины пластинки и линзы интерференционные эффекты, связанные с отражением от других поверхностей, не возникают.) При нормальном падении света интерференционная картина имеет вид концентрических окружностей (кольца Ньютона). Каждая из таких интерференционных окружностей возникает в результате отражения от участков воздушного зазора с одинаковой толщиной (вследствие чего их называют **интерференционными полосами равной толщины**). Определим радиусы колец Ньюто-

на, получающихся при падении света по нормали к пластинке. В этом случае  $\sin i_{\text{п}} = 0$  и оптическая разность хода равна (в воздушном зазоре  $n = 1$ ):

$$\Delta L = 2d + \frac{\lambda}{2}. \quad (29.28)$$

Из треугольника  $CAB$  имеем ( $R \gg d$ )

$$r_m^2 = R^2 - (R - d)^2 = 2Rd - d^2 \approx 2Rd, \quad (29.29)$$

или

$$d = \frac{r_m^2}{2R}. \quad (29.30)$$

Таким образом, для светлых колец (максимумы)

$$\Delta L = \frac{r_m^2}{R} - \frac{\lambda}{2} = m\lambda, \quad m = 0, 1, 2, \dots \quad (29.31)$$

Отсюда следует выражение для радиусов светлых колец:

$$r_m = \sqrt{R\lambda \left( m + \frac{1}{2} \right)}, \quad m = 0, 1, 2, \dots \quad (29.32)$$

Радиусы темных колец Ньютона получаются равными

$$r_m = \sqrt{R\lambda m}, \quad m = 0, 1, 2, \dots \quad (29.33)$$

В точке касания пластинки и линзы, т. е. при  $m = 0, r_0 = 0$ , наблюдается минимум интенсивности (темное пятно), обусловленный изменением фазы колебаний на  $\pi$  при отражении световой волны от пластинки.

**Задача 29.16.** В отраженном свете с длиной волны  $\lambda = 546$  нм были измерены радиусы двух светлых колец, оказавшиеся равными 0,161 и 0,284 см. Было подсчитано, что между этими кольцами расположено 19 других светлых колец. Чему равен радиус  $R$  кривизны линзы?

**Решение.** Пусть меньшее измеренное кольцо Ньютона имеет порядок интерференции  $m$ , тогда порядок большего равен  $m + 20$ . Используем (29.32):

$$r_m = \sqrt{R\lambda \left( m + \frac{1}{2} \right)}, \quad r_{m+20} = \sqrt{R\lambda \left( m + 20 + \frac{1}{2} \right)}.$$

Возведем эти равенства в квадрат и разделим одно на другое:

$$\frac{r_m^2}{r_{m+20}^2} = \frac{m + 0,5}{m + 20,5}.$$

Отсюда находим порядок  $m$  меньшего из упомянутых колец Ньютона:

$$m = \frac{20,5r_m^2 - 0,5r_{m+20}^2}{r_{m+20}^2 - r_m^2} = \frac{20,5 \times 0,161^2 - 0,5 \times 0,284^2}{0,284^2 - 0,161^2} = 9.$$

Искомый радиус кривизны равен:

$$R = \frac{r_9^2}{9,5\lambda} = \frac{(1,61 \cdot 10^{-3})^2}{9,5 \times 546 \cdot 10^{-9}} = 0,5 \text{ м.} \quad \blacksquare$$

## 29.7 Интерферометры

Явление интерференции лежит в основе принципа действия точных измерительных устройств, называемых *интерферометрами*. В этих устройствах пучок электромагнитного излучения (света или радиоволн) каким-то образом разделяется на два или более когерентных пучка, которые после прохождения путей разной оптической длины снова соединяются, формируя интерференционную картину. Небольшие различия в оптической длине пути разных лучей фиксируются по сдвигу интерференционных полос. Интерферометры используются в станкостроении для точного измерения длин, для проверки качества оптических поверхностей, ну и в науке, разумеется. Существует множество вариантов подобных устройств, носящих имена их изобретателей. Наверное, самым знаменитым из них является интерферометр Майкельсона—Морли, с помощью которого была открыта независимость скорости света от движения источника и наблюдателя, что привело к созданию специальной теории относительности. В главе 9 т. 1 была объяснена принципиальная схема опыта Майкельсона.

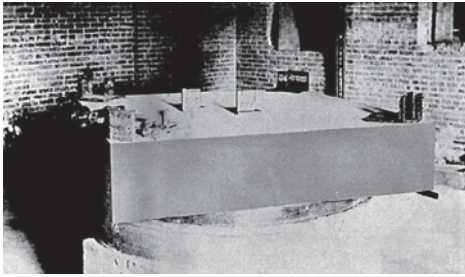
Интерферометр Майкельсона—Морли показан на фотографии рис. 29.9. Это установка, смонтированная в подвале физической лаборатории на каменной плите с длиной стороны всего около полутора метров. Плита плавала в жидкой ртути, опыты проводились по ночам, и все эти ухищрения были направлены на уменьшение вибрации (в частности, от проезжающих мимо здания повозок), которая могла бы помешать зафиксировать искомый сигнал. Чем больше путь, проходимый лучами в интерферометре, тем легче обнаружить разность их хода. Используя многократное отражение луча, Майкельсон и Морли (1887) сумели довести длину пути луча до 11 м.

В феврале 2016 г. международная группа ученых, работавших в США на двух лазерных интерферометрах обсерватории гравитационных волн *LIGO*<sup>2</sup>, сообщила о долгожданном открытии гравитационных волн. Эти интерферометры повторяют в гигантском масштабе схему интерферометра Майкельсона—Морли: длина их плеч составляет 4 км (рис. 29.10). Такие расстояния вместе с использованием лазерной техники позволили на много порядков повысить чувствительность интерферометров. Невероятно, но факт: исследователи сообщили, что «пойманная» ими гравитационная волна изменила длину плеча интерферометра на  $2,5 \cdot 10^{-17}$  м, что в 35 раз меньше радиуса протона. Интерферометры *LIGO* расположены в Ливингстоне (штат Луизиана) и на тихоокеанском побережье в Хэнфорде (штат Вашингтон) на расстоянии 3000 км друг от друга. Идентичный сигнал от гравитационной волны зафиксирован на обоих интерферометрах с разницей во времени 7 мс — это позволило определить угол между направлением на источник и линией, соединяющей детекторы. В июне 2016 г. группа *LIGO* сообщила об обнаружении еще одного сигнала от гравитационной волны.

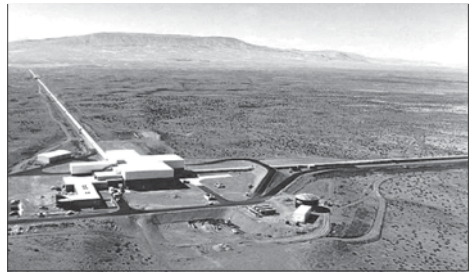
Однако и до открытия гравитационных волн интерферометры уже совершили революцию в астрономии. Увеличение размеров одиночных антенн привело к невероятному росту их чувствительности, но низкая разрешающая способность по углам не позволяла использовать все их преимущества. Кроме того, дальнейшее увеличение размеров конструкций упиралось в прочность материалов

---

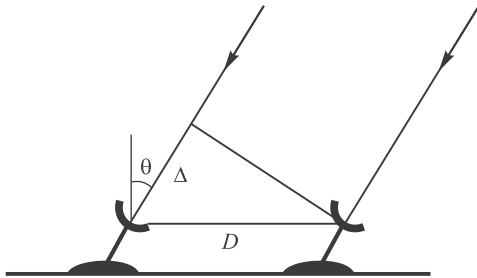
<sup>2</sup>Аббревиатура от англ. *Laser Interferometer Gravitational-Wave Observatory* — лазерно-интерферометрическая гравитационно-волновая обсерватория.



**Рис. 29.9.** Интерферометр Майкельсона—Морли (1887)



**Рис. 29.10.** LIGO — лазерный интерферометр обсерватории гравитационных волн (2016)



**Рис. 29.11.** Схема работы двух телескопов в режиме интерферометра



**Рис. 29.12.** Радиointерферометр VLA, США

и стоимость установок. Выход нашли в соединении данных с нескольких телескопов с помощью электроники и программного обеспечения — как говорят, телескопы работают в режиме интерферометра. Так, в обсерватории Кека на Гавайях построены два телескопа с зеркалами диаметром 10 м. Расстояние между телескопами составляет 85 м, их работа в режиме интерферометра позволяет добиться такого углового разрешения, которое можно было бы получить лишь с зеркалом такого же диаметра.

Принцип работы телескопов в режиме интерферометра пояснен на рис. 29.11. Если на две антенны падает под углом  $\theta$  излучение из космоса и расстояние между приемными антеннами (зеркалами) равно  $D$ , то разность хода лучей, попадающих на соседние антенны, составляет  $\Delta = D \sin \theta$ . При  $\theta = 0$  обе волны находятся в фазе и усиливают друг друга. Первый минимум получается, когда разность хода составляет половину длины волны излучения:  $\Delta = \lambda/2$ , чему соответствуют углы  $\pm\theta_{\min}$ , где

$$\theta_{\min} = \arcsin \left( \frac{\lambda}{2D} \right). \quad (29.34)$$

Таким образом, угловая ширина центрального лепестка диаграммы направленности соединенных антенн (зеркал) равна  $\theta_c = 2\theta_{\min}$ . Она и дает оценку разрешающей способности телескопа. При  $\lambda \ll D$  имеем

$$\theta_c = \frac{\lambda}{D}. \quad (29.35)$$

Таким образом, угловая разрешающая способность повышается с увеличением расстояния между антеннами и при использовании максимально коротких волн.

В штате Нью-Мехико работает радиоинтерферометр *VLA* (англ. *Very Large Array* — Сверхбольшая антенная решетка) Национальной радиоастрономической обсерватории США (рис. 29.12). Он представляет собой 27 радиотелескопов с антеннами диаметром 25 м, расположенных в виде буквы *Y* с длиной плеч более 22 км. Используются длины волн в диапазоне  $\lambda = (0,6-410)$  см. Весь комплекс эквивалентен антенне диаметром  $D = 36$  км.

Но это не предел: совсем уже не выглядят фантастикой радиоинтерферометры, антенны которых расположены на разных континентах или даже в космосе.

Значительным достижением радиоастрономии стал проект «Радиоастрон» с ведущим участием России (НПО им. С. А. Лавочкина и Физический институт им. П. Н. Лебедева РАН). Сердцем проекта является крупнейший в мире космический радиотелескоп с диаметром зеркала 10 м, запущенный в 2011 г. на вытянутую эллиптическую орбиту с максимальным удалением от Земли до 350 000 км. Космический телескоп работает в режиме интерферометра с рядом наземных телескопов. Огромное плечо этого радиоинтерферометра, сравнимое с расстоянием до Луны, позволяет получить самое высокое угловое разрешение за всю историю астрономии. Используя формулу (29.35), мы можем оценить ширину интерференционного лепестка — угловую разрешающую способность «Радиоастрона» — при работе на самом длинном плече на длине волны  $\lambda = 92$  см:

$$\theta_c = \frac{0,92}{0,35 \cdot 10^9} = 2,63 \cdot 10^{-9} \text{ рад} = 540 \text{ угловых микросекунд.} \quad (29.36)$$

При использовании другого доступного диапазона с длиной волны  $\lambda = 6,2$  см угловая разрешающая способность составит уже 37 угловых микросекунд.

## Контрольные вопросы

1. Что такое интерференция световых волн?
2. Когерентные волны. Условия когерентности световых волн.
3. Укажите способы наблюдения интерференции света.
4. При помощи зеркал Френеля получили интерференционные полосы, пользуясь красным светом. Как изменится интерференционная картина, если тот же опыт провести в фиолетовом свете?
5. Чем объясняется радужная окраска мыльных пузырей?
6. Почему не возникает интерференционной картины от двух фар удаленного автомобиля?
7. Почему интерференционные полосы, хорошо различимые в тонких пленках, не наблюдаются на толстом куске стекла?
8. Каково различие интерференционных картин, полученных в отраженном и проходящем свете?
9. Свет отражается от стеклянной пластинки, помещенной в жидкость с показателем преломления, превышающим показатель преломления стекла. При отражении от границы раздела каких сред будет происходить потеря полуволны?

## Глава 30

# Дифракция света

Дифракцией называется отклонение волн от прямолинейного распространения при их взаимодействии с препятствием. Дифракция, как и интерференция, является волновым эффектом и может наблюдаться для волн любой природы. Вследствие дифракции волны могут попадать в область геометрической тени: звук слышен за углом дома, радиоволны могут распространяться далеко за пределы прямой видимости антенного передатчика, а в центре тени от освещенного диска наблюдается светлое пятно.

При дифракции (как и при интерференции) происходит перераспределение интенсивности в результате суперпозиции волн. В сущности, между дифракцией и интерференцией нет принципиальных различий: по историческим причинам явления при суперпозиции конечного числа волн называют интерференцией, а при суперпозиции бесконечного числа волн — дифракцией.

Вследствие обширности технических применений особое значение имеет дифракция света. Наблюдения дифракционных явлений привели волновую оптику в первой половине XIX в. к победе над корпускулярной теорией. В пределе малых длин волн законы волновой оптики переходят в законы геометрической оптики. Следовательно, отклонения от законов геометрической оптики наиболее ярко проявляются, когда размеры препятствий (различных неоднородностей) сопоставимы с длиной световой волны.

### 30.1 Принцип Гюйгенса—Френеля

Изложенный в разделе 28.1 принцип Гюйгенса имеет характер геометрического правила. Согласно принципу Гюйгенса, каждую точку волнового фронта можно рассматривать как самостоятельный источник колебаний, а результат действия вторичных волн может быть найден построением поверхности, огибающей эти вторичные волны. Французский физик О. Френель дополнил этот принцип, предложив рассматривать волновое возмущение в любой точке пространства как результат *интерференции вторичных волн* от фиктивных источников, в качестве которых в каждый момент времени выступают все точки фронта распространяющейся в пространстве волны. Эти фиктивные источники когерентны и поэтому могут интерферировать в любой точке пространства, в результате

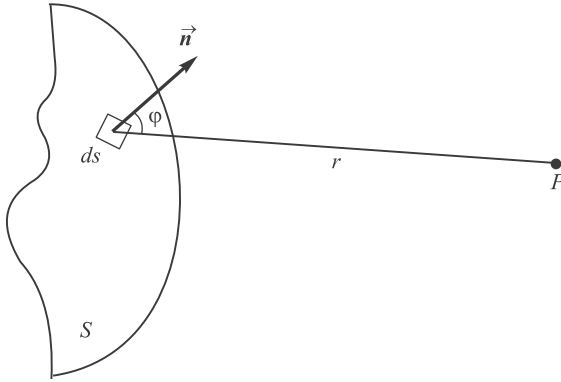


Рис. 30.1. Применение принципа Гюйенса–Френеля к сферической волне

чего элементарные волны могут гасить или усиливать друг друга. Способ Френеля вкладывает более глубокое физическое содержание в принцип Гюйенса. Он также позволяет решить ряд задач, представлявших трудности в рамках первоначального принципа Гюйенса.

Рассмотрим поверхность волнового фронта  $S$  (рис. 30.1). Согласно *принципу Гюйенса–Френеля*

♦ *каждый элемент поверхности волнового фронта служит источником вторичной сферической волны, амплитуда которой пропорциональна площади  $ds$  элемента.* ♦

Амплитуда сферической волны убывает с расстоянием  $r$  от источника как  $1/r$ . Следовательно, от каждого элемента  $ds$  волновой поверхности в точку наблюдения  $P$  приходит световое возмущение

$$dE = K(\varphi) \frac{A_0}{r} \cos(\omega t - kr + \alpha_0) ds, \quad (30.1)$$

где  $A_0 ds$  — амплитуда сферической волны, испускаемой элементом волновой поверхности площадью  $ds$ ; величина  $K(\varphi)$  — коэффициент, который зависит от угла  $\varphi$  между нормалью  $\vec{n}$  к площадке  $ds$  и направлением волнового вектора  $\vec{k}$  (или, что то же самое, направлением, соединяющим элемент поверхности  $ds$  с точкой наблюдения  $P$ ). Результирующее колебание в точке  $P$  представляет суперпозицию элементарных колебаний  $dE$ , причем интеграл берется по всей волновой поверхности  $S$ :

$$E = \iint_S K(\varphi) \frac{A_0}{r} \cos(\omega t - kr + \alpha_0) ds. \quad (30.2)$$

Это соотношение представляет аналитическое выражение принципа Гюйенса–Френеля. В общем случае расчет интерференции вторичных волн довольно сложен и громоздок, однако в ряде случаев нахождение амплитуды результирующего колебания оказывается возможным с помощью алгебраического или геометрического суммирования.

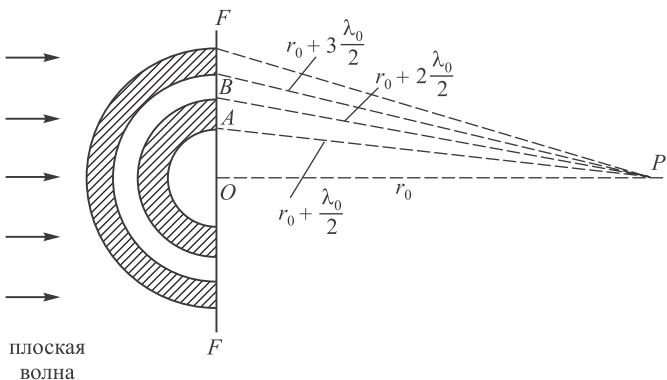
## 30.2 Метод зон Френеля

Принцип Гюйгенса—Френеля в рамках волновой теории света позволяет объяснить прямолинейное распространение света. Определим амплитуду световой волны в произвольной точке  $P$ , используя *метод зон Френеля*. Рассмотрим сначала случай падающей плоской волны (рис. 30.2).

Пусть плоский фронт волны  $F$ , распространяющейся от расположенного в бесконечности источника света, в некоторый момент времени находится на расстоянии  $OP = r_0$  от точки наблюдения  $P$ . Во всех точках фронта волны, согласно принципу Гюйгенса—Френеля, возникают элементарные сферические волны, которые распространяются по всем направлениям и через некоторое время достигают точки наблюдения  $P$ . Результирующая амплитуда колебаний в этой точке определяется векторной суммой амплитуд всех вторичных волн. Колебания всех точек волнового фронта  $F$  происходят в одной фазе. В то же время все точки фронта  $F$  находятся от точки  $P$  на различных расстояниях. Для определения результирующей амплитуды всех волн Френель предложил метод разбиения волновой поверхности на кольцевые зоны — *зоны Френеля*.

Взяв точку  $P$  в качестве центра, построим ряд концентрических сфер, радиусы которых начинаются с  $r_0$  и увеличиваются каждый раз на половину длины волны  $\lambda/2$ . При пересечении с плоским фронтом волны  $F$  эти сферы дадут концентрические окружности. Таким образом, на фронте волны появятся кольцевые зоны (зоны Френеля) с радиусами  $\rho_1, \rho_2, \rho_3$  и т. д. Определим радиусы зон Френеля:  $OA = \rho_1$ ,  $OA^2 = AP^2 - OP^2$ , т. е.

$$\rho_1^2 = \left(r_0 + \frac{\lambda}{2}\right)^2 - r_0^2 = r_0\lambda + \frac{\lambda^2}{4} \approx r_0\lambda \quad (\lambda \ll r_0). \quad (30.3)$$



**Рис. 30.2.** Применение принципа Гюйгенса—Френеля к плоской волне: зоны Френеля на поверхности плоского волнового фронта  $F$  представляют собой концентрические кольца (для наглядности их изображение развернуто на  $90^\circ$  и они показаны такими, какими выглядят из точки  $P$ ; на самом деле зоны Френеля расположены в плоскости, перпендикулярной линии  $OP$ )



Аналогично находим:

$$\begin{aligned}\rho_2^2 &= \left(r_0 + 2\frac{\lambda}{2}\right)^2 - r_0^2 = 2r_0\lambda + \lambda^2 \approx 2r_0\lambda, \\ \rho_3^2 &= \left(r_0 + 3\frac{\lambda}{2}\right)^2 - r_0^2 = 3r_0\lambda + 9\frac{\lambda^2}{4} \approx 3r_0\lambda, \\ \dots & \quad \dots \\ \rho_n^2 &\approx nr_0\lambda, \quad n = 1, 2, 3, \dots\end{aligned}\quad (30.4)$$

Для оценки амплитуд колебаний определим площади зон Френеля. Первая зона (круг):

$$S_1 = \pi\rho_1^2 = \pi r_0\lambda; \quad (30.5)$$

вторая зона (кольцо):

$$S_2 = \pi\rho_2^2 - \pi\rho_1^2 = \pi r_0\lambda; \quad (30.6)$$

третья и последующие зоны (кольца):

$$\begin{aligned}S_3 &= \pi\rho_3^2 - \pi\rho_2^2 = \pi r_0\lambda, \\ \dots & \quad \dots \\ S_n &= \pi r_0\lambda.\end{aligned}\quad (30.7)$$

Таким образом, площади зон Френеля примерно одинаковы (при условии  $\lambda \ll r_0$ ).

Подчеркнем, что согласно принципу Гюйгенса—Френеля каждый элемент волновой поверхности служит источником вторичных сферических волн, *амплитуда которых пропорциональна площади этого элемента поверхности*. Важно также понять, что зоны Френеля — это всего лишь удобный математический прием, их расположение не есть свойство самой волны. Оно зависит от точки наблюдения  $P$ : переместившись в какую-либо другую точку, наблюдатель совсем иначе разобьет мысленно тот же волновой фронт на зоны Френеля.

Колебания, возбуждаемые в точке  $P$  двумя соседними зонами, *противоположны по фазе*, так как разность хода лучей от двух соседних зон до точки наблюдения  $P$  равна  $\lambda/2$ . Следовательно, волны приходят в точку  $P$  в противофазе, и поэтому при наложении они взаимно ослабляют друг друга. Таким образом, амплитуда  $A$  результирующего колебания может быть представлена в виде знакопеременного ряда

$$A = A_1 - A_2 + A_3 - A_4 + \dots, \quad (30.8)$$

где  $A_1$  — амплитуда колебаний в точке  $P$ , возбуждаемых действием центральной (первой) зоны Френеля,  $A_2$  — амплитуда колебаний, возбуждаемых второй зоной, и т. д.

Расстояние от  $n$ -й зоны до точки  $P$  медленно растет с номером зоны  $n$ . Угол  $\varphi$  между нормалью к элементам зоны и направлением в точку  $P$  также растет с  $n$ . Следовательно, амплитуда  $A_n$  колебания, возбуждаемого  $n$ -й зоной в точке  $P$ , монотонно убывает с ростом  $n$ . Другими словами, амплитуды колебаний, возбуждаемых в точке  $P$  зонами Френеля, образуют монотонно убывающую последовательность:

$$A_1 > A_2 > \dots > A_{n-1} > A_n > A_{n+1} > \dots \quad (30.9)$$

Вследствие монотонного и медленного убывания  $A_n$  можно приближенно считать, что амплитуда колебаний от зоны с номером  $n$  равна среднему арифметическому от амплитуд колебаний от двух соседних зон Френеля:

$$A_n = \frac{A_{n-1} + A_{n+1}}{2}. \quad (30.10)$$

В формуле (30.8) все амплитуды от четных зон входят с одним знаком, а от нечетных — с другим. Запишем это равенство в следующем виде:

$$A = \frac{A_1}{2} + \left( \frac{A_1}{2} - A_2 + \frac{A_3}{2} \right) + \left( \frac{A_3}{2} - A_4 + \frac{A_5}{2} \right) + \dots \quad (30.11)$$

Вследствие соотношений (30.10) выражения в скобках равны нулю, так что

$$A = \frac{A_1}{2}, \quad (30.12)$$

т. е. результирующая амплитуда, создаваемая в точке наблюдения  $P$  всей поверхностью волнового фронта, равна половине амплитуды, создаваемой одной лишь центральной (первой) зоной Френеля. Таким образом, колебания, вызываемые в точке  $P$  волновой поверхностью  $F$ , имеют такую же амплитуду, как если бы действовала только половина первой (центральной) зоны. Следовательно, свет распространяется как бы в узком канале, сечение которого равно половине первой (центральной) зоны Френеля. Мы снова пришли к прямолинейному распространению плоской волны света.

Если же на пути волны поставить диафрагму с отверстием, оставляющим открытой только центральную (первую) зону Френеля, амплитуда в точке  $P$  будет равна  $A_1$ , т. е. в 2 раза превысит амплитуду, создаваемую всем волновым фронтом. Соответственно интенсивность света в точке  $P$  будет в 4 раза больше, чем при отсутствии преграды между источником света и точкой  $P$ . Удивительно, не так ли? Но чудес в природе не бывает: в других точках плоскости, перпендикулярной линии  $OP$  и проходящей через точку  $P$ , интенсивность света будет ослаблена, а средняя освещенность всего экрана при использовании диафрагмы уменьшится, как и следовало ожидать.

Возможность деления волнового фронта на зоны Френеля подтверждена экспериментально. Колебания от четных и нечетных зон Френеля находятся в противофазе и, следовательно, взаимно ослабляют друг друга. Если поставить на пути световой волны пластинку, которая перекрывала бы все четные или нечетные зоны, то интенсивность света в точке  $P$  резко возрастет. Такая **зонная пластинка** действует подобно собирающей линзе. Подчеркнем еще раз: зоны Френеля — это мысленно выделенные участки поверхности волнового фронта, положение которых зависит от выбранной точки наблюдения  $P$ . При другой точке наблюдения расположение зон Френеля будет иным. Метод зон Френеля — удобный способ решения задач на дифракцию волн на тех или иных препятствиях.

Различают два вида дифракции. Если источник света  $S$  и точка наблюдения  $P$  находятся далеко от препятствия, лучи, падающие на препятствие и идущие в точку  $P$ , образуют практически параллельные пучки. В таком случае говорят о *дифракции в параллельных лучах*, или **дифракции Фраунгофера**. Если же рассматривается дифракционная картина на конечном расстоянии от препятствия, вызвавшего дифракцию, то говорят о *дифракции сферических волн*, или **дифракции Френеля**.

### 30.3 Дифракция на круглом отверстии и диске

Дифракционные явления можно наблюдать при прохождении света через малое отверстие или вблизи экрана (диска). Плоская световая волна падает на малое круглое отверстие радиусом  $a$  (рис. 30.3). Плоский фронт, совпадающий с отверстием, можно рассматривать как совокупность фиктивных источников, испускающих когерентные волны, которые в точке наблюдения  $P$  будут интерферировать. Разобьем площадь отверстия на ряд кольцевых зон Френеля, для чего из точки  $P$  проведем ряд сфер с радиусами  $r_1 = r_0 + \lambda/2$ ,  $r_2 = r_0 + 2\lambda/2$  и т. д. Если число зон, которые укладываются в отверстии, четно, то в точке  $P$  будет темное пятно. Действительно, результирующая амплитуда колебаний при четном числе зон  $n = 2m$  равна:

$$A = \frac{A_1}{2} + \left( \frac{A_1}{2} - A_2 + \frac{A_3}{2} \right) + \left( \frac{A_3}{2} - A_4 + \frac{A_5}{2} \right) + \dots + \left( \frac{A_{2m-1}}{2} - A_{2m} + \frac{A_{2m+1}}{2} \right) - \frac{A_{2m+1}}{2} = \frac{A_1}{2} - \frac{A_{2m+1}}{2}. \quad (30.13)$$

При небольших отверстиях (небольших  $m$ ) амплитуды мало отличаются друг от друга и  $A \approx 0$ . При нечетном числе зон  $n = 2m + 1$  аналогичные рассуждения приводят к выражению

$$A = \frac{A_1}{2} + \frac{A_{2m+1}}{2} \approx A_1, \quad (30.14)$$

т. е. в точке наблюдения будет светлое пятно. Число  $n$  зон Френеля, расположенных в отверстии, зависит от расстояния  $r_0$  и длины волны  $\lambda$ . Приравнявая внешний радиус  $\rho_n$  последней зоны Френеля радиусу  $a$  отверстия, получаем из (30.4) соотношение

$$\rho_n^2 = nr_0\lambda = a^2, \quad (30.15)$$

откуда следует выражение для числа зон Френеля, помещающихся в отверстии:

$$n = \frac{a^2}{r_0\lambda}. \quad (30.16)$$

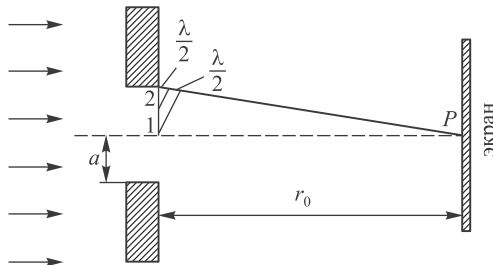


Рис. 30.3. Дифракция света на круглом отверстии

Таким образом, при данном радиусе отверстия  $a$  и длине волны  $\lambda$  падающего света число зон  $n$  является функцией расстояния  $r_0$  между отверстием и точкой наблюдения.

Расчет амплитуды результирующих колебаний, пришедших в другие точки экрана, более сложен. Из соображений симметрии следует, что интерференционная картина на экране вокруг центрального светлого (или темного) пятна (в зависимости от четности  $n$ ) должна иметь вид чередующихся светлых и темных колец с центрами в точке  $P$ . Интенсивность максимумов должна убывать при удалении от точки  $P$ .

Если источник света расположен перед отверстием на не слишком большом расстоянии  $r$  до него, то расчет зон Френеля слегка усложняется: зоны проводятся не на плоском, а на сферическом фронте. Приведем без вывода выражение для радиуса зон Френеля в этом случае:

$$\rho_n^2 = n\lambda \frac{rr_0}{r + r_0}. \quad (30.17)$$

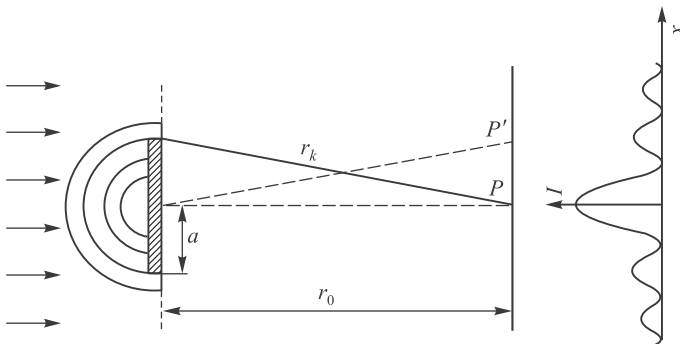
При  $r \rightarrow \infty$  мы получаем случай плоской волны (30.4).

Поместим теперь между падающей плоской волной и точкой наблюдения  $P$  непрозрачный диск радиусом  $a$  (рис. 30.4). Если диск закроет  $n$  первых зон Френеля, то амплитуда в точке  $P$  будет равна:

$$\begin{aligned} A &= A_{n+1} - A_{n+2} + A_{n+3} - \dots = \\ &= \frac{A_{n+1}}{2} + \left( \frac{A_{n+1}}{2} - A_{n+2} + \frac{A_{n+3}}{2} \right) + \dots \end{aligned} \quad (30.18)$$

Выражения в скобках можно положить равными нулю, т. е.

$$A = \frac{A_{n+1}}{2}. \quad (30.19)$$



**Рис. 30.4.** Дифракция света на непрозрачном диске: справа показано распределение освещенности экрана в зависимости от расстояния  $x$  от центра экрана. Светлое пятно в центре (максимальное значение интенсивности света  $I$ ) сменяется чередующимися минимумами и максимумами, образующими светлые и темные концентрические кольца

Таким образом, за небольшим ( $A_{n+1} \approx A_1$ ) непрозрачным диском в центре экрана будет светлое пятно. Для точки  $P'$ , смещенной относительно точки  $P$  в любом радиальном направлении, диск будет перекрывать часть  $(n + 1)$ -й зоны Френеля, одновременно откроется часть зоны  $n$ . Это вызовет уменьшение интенсивности. При некотором положении точки  $P'$  интенсивность достигает минимума. Следовательно, в случае непрозрачного круглого диска дифракционная картина имеет вид светлого центрального пятна и чередующихся темных и светлых концентрических колец (см. рис. 30.4). Светлое пятно в центре геометрической тени было предсказано С. Пуассоном в 1818 г. и выдвинуто в качестве опровержения волновой теории света. Однако Д. Араго провел соответствующий опыт и показал, что выводы Пуассона соответствуют действительности и лишь подтверждают волновую теорию и ее предсказания, вытекающие из метода зон Френеля.

**Задача 30.17.** На диафрагму с круглым отверстием диаметром  $d = 4$  мм падает по нормали плоская волна света ( $\lambda = 0,5$  мкм). Точка наблюдения находится на расстоянии  $b = 1$  м на оси отверстия. Сколько зон Френеля укладывается в отверстие? Светлое или темное пятно получится в центре дифракционной картины?

**Решение.** Используем (30.16) при  $a = d/2$  и  $r_0 = b$ :

$$n = \frac{d^2}{4b\lambda} = \frac{(4 \cdot 10^{-3})^2}{4 \times 1 \times 0,5 \cdot 10^{-6}} = 8.$$

В центре картины будет темное пятно. ■

**Задача 30.18.** Точечный источник света ( $\lambda = 0,5$  мкм) расположен на расстоянии  $l = 1$  м на оси диафрагмы с отверстием радиусом  $a = 1$  мм. За отверстием помещают экран. При каком расстоянии от отверстия до экрана для центра дифракционной картины будут открыты  $n = 3$  зоны Френеля? Как изменится освещенность в центре, если убрать диафрагму?

**Решение.** Используем формулу (30.17):

$$a^2 = n\lambda \frac{lr_0}{l + r_0}.$$

Отсюда находим:

$$r_0 = \frac{la^2}{n\lambda l - a^2} = \frac{1 \times (10^{-3})^2}{3 \times 0,5 \cdot 10^{-6} \times 1 - (10^{-3})^2} = 2 \text{ м.}$$

В центре дифракционной картины будет светлое пятно, а амплитуда колебаний будет равна  $A \approx A_1$ . Если диафрагму убрать, то амплитуда станет равной  $A_1/2$ , т. е. освещенность в этой точке уменьшится в четыре раза. ■

## 30.4 Дифракция Фраунгофера от щели

Пусть на бесконечно длинную щель падает плоская световая волна. В соответствии с принципом Гюйгенса—Френеля освещенную щель можно рассматривать как множество точечных когерентных источников света. Поместим за щелью экран, расстояние до которого достаточно велико по сравнению с шириной щели. Это условие означает, что в данную точку  $P$  экрана попадет параллельный пучок

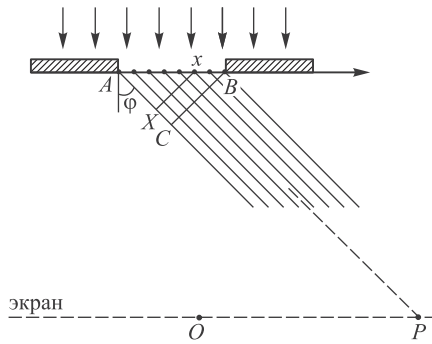


Рис. 30.5. Дифракция Фраунгофера от щели

лучей, отклонившийся на угол  $\varphi$  (рис. 30.5). Разность хода  $AC = \Delta$  крайних лучей из этого пучка определяется из треугольника  $ABC$  ( $\angle ABC = \varphi$ ):

$$\Delta = a \sin \varphi, \quad (30.20)$$

где  $a = AB$  — ширина щели. Если при наблюдении из точки  $P$  в щели помещается четное число зон Френеля ( $\Delta = 2m\lambda/2$ ), то их вклады взаимно погасятся и в точке  $P$  будет наблюдаться минимум интенсивности света. Таким образом, уравнение

$$\sin \varphi_{\min, m} = 2m \frac{\lambda}{2a}, \quad m = 1, 2, \dots \quad (30.21)$$

дает нам **условие дифракционных минимумов**, где угол  $\varphi_{\min, m}$  — направление на минимум порядка  $m$ .

Если разность хода крайних лучей равна нечетному числу полуволн, т. е.  $\Delta = (2m + 1)\lambda/2$ , то при наблюдении из точки  $P$  в щели помещается нечетное число зон Френеля. Каждая зона гасит соседнюю, а оставшаяся последняя посылает свет в направлении  $\varphi_{\max, m}$  и образует максимум. Поэтому **условие максимумов** имеет вид

$$\sin \varphi_{\max, m} = (2m + 1) \frac{\lambda}{2a}, \quad m = 1, 2, \dots \quad (30.22)$$

Выражения (30.21) и (30.22) имеют приближенный характер, поскольку при их выводе сделаны определенные предположения (хотя и разумные) о вкладе световых потоков от отдельных зон.

Что же касается центральной точки  $O$  экрана, то в нее попадает пучок неотклоненных лучей, ортогональных щели. Все они имеют одинаковую фазу, т. е. должны усиливать друг друга. Поэтому в условии минимумов (30.21) исключено значение  $m = 0$ , соответствующее точке  $O$ . Это же значение исключено и в условии максимумов (30.22): если бы угол  $\varphi_{\max, 0} = \arcsin(\lambda/(2a))$  соответствовал какому-либо максимуму, то непонятно, что бы мы наблюдали

между ним и *центральный максимум*  $\varphi_{\max,c} = 0$ , так как между ними, как следует из (30.22), не должно быть минимумов интенсивности, поскольку  $\varphi_{\min,1} = \arcsin(\lambda/a) > \varphi_{\max,0}$ .

После этих качественных соображений изучим дифракционную картину более подробно и получим выражения, позволяющие сравнить интенсивности света в максимумах различных порядков. Результирующее колебание в некоторой точке  $P$  экрана представляет собой суперпозицию колебаний, распространяющихся от всей поверхности щели. В случае дифракции Фраунгофера расстояние от экрана до решетки можно считать приблизительно постоянным при малых углах  $\varphi$ . Коэффициент  $K(\varphi)$  в общей формуле (30.2) также можно считать постоянным, если мы ограничимся рассмотрением не слишком больших углов дифракции  $\varphi$ . Обозначим через  $A_0$  сумму амплитуд колебаний, возбуждаемых щелью в центральной точке  $O$  экрана. Тогда амплитуда колебаний, возбуждаемых элементом щели длиной  $dx$ , будет равна:

$$dA = \frac{A_0}{a} dx. \quad (30.23)$$

Такой же будет амплитуда колебаний, возбуждаемых этим же элементом в любой другой точке  $P$ . Однако если этот элемент находится в точке с координатой  $x$  (начало координат мы поместим в крайнюю точку  $A$  щели), то вторичная волна, дошедшая от него до точки  $P$ , будет опережать по фазе колебание, дошедшее до точки  $P$  от точки  $A$ . Разность фаз между рассматриваемыми колебаниями образуется на пути  $AX = x \sin \varphi$  (см. рис. 30.5). Если начальную фазу колебания, возбуждаемого в точке  $P$  элементарной площадкой, расположенной в точке  $A$ , считать равной нулю, то начальная фаза колебания, возбуждаемого площадкой с координатой  $x$ , будет равна:

$$\alpha_{0,x} = k \cdot AX = kx \sin \varphi, \quad (30.24)$$

где  $k = 2\pi/\lambda$  — волновое число световой волны. Таким образом, учитывая (30.23) и (30.24), находим колебание, возбуждаемое в точке  $P$  элементом щели с координатой  $x$ :

$$dE_\varphi = \frac{A_0}{a} \cos(\omega t + kx \sin \varphi) dx. \quad (30.25)$$

Проинтегрируем это соотношение по всей ширине щели ( $0 \leq x \leq a$ ) и получим результирующее колебание, возбуждаемое в точке  $P$ :

$$\begin{aligned} E_\varphi &= \frac{A_0}{a} \int_0^a \cos(\omega t + kx \sin \varphi) dx = \frac{A_0}{a} \cdot \frac{1}{k \sin \varphi} \sin(\omega t + kx \sin \varphi) \Big|_0^a = \\ &= \frac{A_0}{a} \cdot \frac{\sin(\omega t + ka \sin \varphi) - \sin \omega t}{k \sin \varphi} = \\ &= A_0 \frac{\sin\left(\frac{ka}{2} \sin \varphi\right)}{\frac{ka}{2} \sin \varphi} \cos\left(\omega t + \frac{ka}{2} \sin \varphi\right). \end{aligned} \quad (30.26)$$

Учитывая связь  $k = 2\pi/\lambda$  волнового вектора с длиной волны, приходим к выражению для амплитуды результирующего колебания:

$$A_1(\varphi) = A_0 \left| \frac{\sin\left(\frac{\pi a}{\lambda} \sin \varphi\right)}{\frac{\pi a}{\lambda} \sin \varphi} \right|. \quad (30.27)$$

Для точки  $O'$ , лежащей против центра щели, угол  $\varphi = 0$  и  $A_1 = A_0$ . Этот результат следует, как мы видели, и из физических рассуждений. Получим положение других максимумов. Для этого представим результирующую амплитуду в виде

$$A_1 = A_0 \left| \frac{\sin \psi}{\psi} \right|, \quad \psi = \pi \frac{a}{\lambda} \sin \varphi. \quad (30.28)$$

Амплитуда имеет максимум при выполнении условия

$$\frac{d}{d\psi} \frac{\sin \psi}{\psi} = 0 = \frac{\cos \psi}{\psi} - \frac{\sin \psi}{\psi^2}, \quad (30.29)$$

или

$$\operatorname{tg} \psi = \psi. \quad (30.30)$$

Очевидное решение  $\psi = 0$  соответствует центральному максимуму. Следующие за ним корни уравнения (30.30), которое может быть решено только численно, равны  $\psi_1 = \pm 4,4934$ ,  $\psi_2 = \pm 7,7253$ ,  $\psi_3 = \pm 10,904$ . Отсюда находим условие первого максимума:

$$\sin \varphi_1 = \pm \frac{4,49341}{\pi} \frac{\lambda}{a} = \pm 1,43 \frac{\lambda}{a}. \quad (30.31)$$

Из приближенного выражения (30.22) при  $m = 1$  следует коэффициент 1,5 вместо правильного 1,43, что всего на 5% больше. Для других максимумов согласие с приближенной формулой становится еще лучше:

$$\sin \varphi_2 = \pm \frac{7,7253}{\pi} \frac{\lambda}{a} = \pm 2,46 \frac{\lambda}{a}, \quad \sin \varphi_3 = \pm \frac{10,904}{\pi} \frac{\lambda}{a} = \pm 3,47 \frac{\lambda}{a}. \quad (30.32)$$

«Подозрительная» же точка  $\sin \varphi = \lambda/(2a)$  вообще не соответствует экстремуму амплитуды (см. (30.27)).

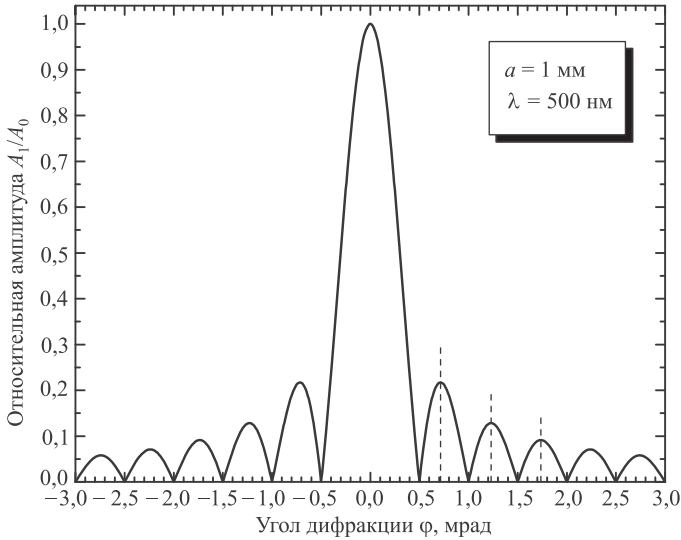
На рисунке 30.6 показана рассчитанная по формуле (30.28) зависимость от угла относительной амплитуды  $A_1/A_0$  для волны длиной  $\lambda = 500$  нм и щели шириной  $a = 1$  мм.

При углах  $\varphi$ , удовлетворяющих условию  $a \sin \varphi = \pm m\lambda$  ( $m = 1, 2, 3, \dots$ ), амплитуда  $A_\varphi$  равна нулю. Это условие определяет положение минимумов, как и было получено выше. Для численного примера с  $\lambda = 500$  нм и  $a = 1$  мм, представленного на рис. 30.6, соответствующие минимумам углы равны  $\varphi_m = 0,5m$  мрад.

Интенсивность света пропорциональна квадрату амплитуды. Следовательно,

$$I_1(\varphi) = I_0 \frac{\sin^2 \psi}{\psi^2}, \quad \psi = \pi \frac{a}{\lambda} \sin \varphi, \quad (30.33)$$





**Рис. 30.6.** Относительная амплитуда  $A_1/A_0$  при дифракции света на одной щели; вертикальные пунктирные линии показывают положения максимумов, рассчитанные по формулам (30.31) и (30.32)

где  $I_0$  — интенсивность в середине дифракционной картины,  $I_1(\varphi)$  — интенсивность в точке  $P$ , положение которой определяется углом  $\varphi$ . Подставляя сюда  $\psi = 4,49341$ , находим интенсивность  $I_1$  в первом максимуме:  $I_1 = 0,047 I_0$ . Иначе, интенсивность в первом максимуме почти в 20 раз меньше, чем в центральном. Интенсивность в других максимумах будет еще меньше.

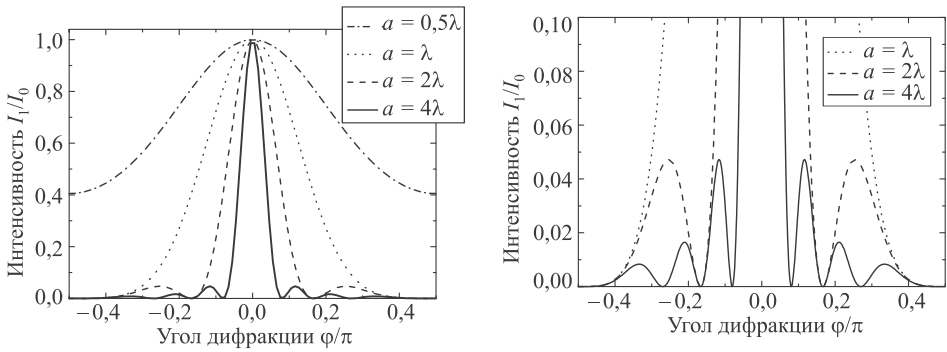
Таким образом, центральный максимум дает главное изображение щели. В качестве меры его ширины можно принять расстояние между минимумами слева и справа от него. Используя условие первых минимумов  $\sin \varphi = \pm \lambda/a$  и учитывая, что при малых углах  $\sin \varphi \sim \varphi$ , получаем что минимумы видны под углами  $\varphi_{\pm} = \pm \lambda/a$ . Поэтому угловой размер центрального минимума равен:

$$\Delta\varphi = \varphi_+ - \varphi_- = \frac{2\lambda}{a}. \quad (30.34)$$

Аналогичные формулы для отверстий другой формы различаются лишь числовым коэффициентом. Отсюда следует общий вывод для любых оптических приборов. Если с помощью оптического прибора (микроскоп, подзорная труба и т. п.) пытаются разглядеть два предмета, угловое расстояние между которыми равно  $\Delta\varphi$ , то это удастся сделать, если

$$\Delta\varphi > \frac{\lambda}{a}. \quad (30.35)$$

Под  $a$  здесь надо понимать линейный размер отверстия прибора — его объектива. Иначе изображения предметов (их центральные максимумы) попадут в одну



**Рис. 30.7.** Слева — распределение интенсивности при дифракции на щелях разной ширины; справа — тот же график в увеличенном по вертикали масштабе

точку, и предметы будет невозможно различить. Для повышения *разрешающей способности прибора*, которая тем выше, чем меньше ширина центрального максимума (см. (30.34)), надо либо увеличить диаметр  $a$  объектива, либо использовать возможно более короткие волны. Последнее реализуется в электронных микроскопах.

**Задача 30.19.** *Может ли орел с высоты  $h = 1$  км разглядеть на земле мышонка размером  $l = 2$  см?*

**Решение.** Угловой размер мышонка  $\Delta\varphi = l/h$ . Чтобы орел мог разглядеть его в видимом диапазоне света (примем  $\lambda = 500$  нм), диаметр  $a$  его зрачка должен удовлетворять соотношению (30.35):

$$a > \frac{\lambda}{\Delta\varphi} = \frac{\lambda h}{l} = \frac{5 \cdot 10^{-7} \times 10^3}{2 \cdot 10^{-2}} = 2,5 \cdot 10^{-2} \text{ м} = 2,5 \text{ см.} \quad (30.36)$$

Ответ на вопрос задачи теперь очевиден, поскольку таких огромных глаз у птиц не бывает. ■

На рисунке 30.7 показана дифракционная картина для щелей разной ширины. Видно, как расширяется дифракционный пик при уменьшении размера щели. С дифракционным уширением светового пучка мы еще встретимся в главе 35, посвященной волновым свойствам микрочастиц.

## 30.5 Дифракционная решетка

Широкое распространение в научном эксперименте и технике получили *дифракционные решетки*, которые представляют собой множество параллельных, расположенных на равных расстояниях одинаковых щелей, разделенных равными по ширине непрозрачными промежутками. Дифракционные решетки изготавливаются с помощью делительной машины, наносящей штрихи (царапины) на стекле или другом прозрачном материале. Там, где проведены царапины, материал становится непрозрачным, а промежутки между ними остаются прозрачными и фактически выполняют роль щелей.

Суть явления состоит в том, что если длина пространственной когерентности намного больше ширины дифракционной решетки, то ее штрихи разбивают световой поток на когерентные пучки. Каждая щель между штрихами порождает свою дифракционную картину (этот процесс рассмотрен в предыдущем разделе). Кроме того, пучки от щелей интерферируют между собой, подобно тому, как интерферировали отдельные световые лучи в опытах Юнга. Вследствие такой **многолучевой интерференции** когерентных пучков дифракционные картины от разных щелей не просто складываются — происходит перераспределение интенсивности светового потока в пространстве. Положение максимумов полу-чающейся интерференционной картины зависит от длины волны, и поэтому дифракционная решетка раскладывает белый свет в цветной спектр, подобно ньютоновской призме<sup>1</sup>.

### Дифракция от двух щелей

Рассмотрим сначала дифракцию света от решетки с двумя щелями. Пусть  $a$  — ширина щели, а  $b$  — ширина непрозрачного промежутка (рис. 30.8). Расстояние между серединами соседних щелей называется **периодом решетки**:

$$d = a + b. \quad (30.37)$$

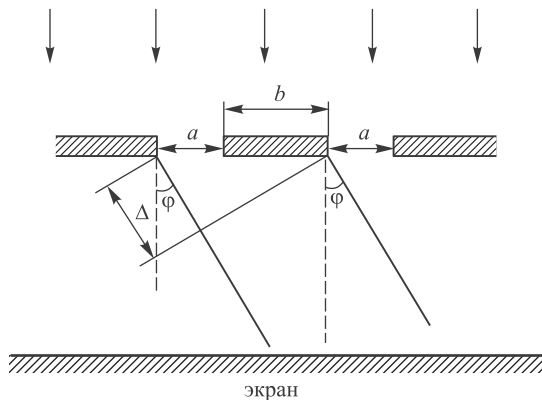


Рис. 30.8. Дифракция от двух щелей

<sup>1</sup>Первым свойства дифракционной решетки обнаружил шотландский математик и астроном Джеймс Грегори (1638–1675). В его опытах роль дифракционной решетки выполняли птичьи перья — как можно более тонкие и желательнее белые, как он сам их описал. Помещая перья перед небольшим отверстием, через которое солнечный свет проникал в темную комнату, Грегори наблюдал на листе бумаги или на стене напротив образованный светом рисунок из кругов и овалов. Центральный круг напротив отверстия был белым, а все остальные были окрашены в разные цвета. Грегори написал об этом своему коллеге-математику в мае 1673 г., через год после опытов Ньютона по разложению белого света в цветной спектр призмой. Он просил сообщить Ньютону о своем «небольшом эксперименте», который «стоит того, чтобы над ним поразмыслить», если Ньютон еще не знает о таком явлении. Свое письмо Грегори закончил словами, что с радостью услышал бы соображения Ньютона по этому поводу.

При угле дифракции  $\varphi$  разность хода двух параллельных крайних (или любых сходственных) лучей, идущих от соседних щелей, равна:

$$\Delta = d \sin \varphi. \quad (30.38)$$

Если эта разность хода равна нечетному числу полуволин

$$\Delta = (2m + 1) \frac{\lambda}{2}, \quad m = 0, 1, 2, 3, \dots, \quad (30.39)$$

то свет, посылаемый двумя щелями, будет взаимно гаситься. **Условие минимумов интерференции волн от соседних щелей** имеет вид

$$\boxed{d \sin \varphi = \pm(2m + 1) \frac{\lambda}{2}.} \quad (30.40)$$

Эти минимумы называются **дополнительными минимумами**.

Если разность хода равна четному числу полуволин

$$\Delta = 2m \frac{\lambda}{2}, \quad m = 0, 1, 2, 3, \dots, \quad (30.41)$$

то свет, посылаемый каждой щелью, будет взаимно усиливать друг друга. **Условие интерференционных максимумов** с учетом (30.38) имеет вид

$$\boxed{d \sin \varphi = \pm 2m \frac{\lambda}{2} = \pm m \lambda.} \quad (30.42)$$

Это формула для **главных максимумов** дифракционной решетки. Целое число  $m$  называют порядком главного максимума или **порядком спектра**: спектр нулевого порядка соответствует центральному, наиболее интенсивному, максимуму при  $\varphi = 0$ .

Кроме того, в направлениях, куда ни одна из щелей не распространяет свет, он не будет распространяться и при двух щелях, т. е. **главные минимумы** решетки будут наблюдаться в направлениях, определяемых условием (30.21) для одной щели:

$$\boxed{a \sin \varphi = \pm m \lambda, \quad m = 1, 2, 3, \dots} \quad (30.43)$$

Таким образом, дифракция на двух щелях приводит к появлению дополнительных максимумов и минимумов. Кроме того, легко понять, что в главных максимумах световой поток от двух щелей складывается и, стало быть, амплитуда становится в 2 раза больше, чем от одной щели, а интенсивность соответственно увеличивается в 4 раза. Можно также догадаться, что одновременно пик сужается, становится более отчетливым, так как в целом световой поток остается тем же, но лишь перераспределяется в пространстве.

## Дифракция на решетке

Рассмотрим теперь дифракцию на решетке, имеющей  $N$  щелей (современные решетки, применяемые в приборах для спектрального анализа, имеют до 200 000 штрихов и период  $d \approx 0,8$  мкм, т. е. порядка 12 000 штрихов на 1 см).

Условие (30.43) главных минимумов не зависит от числа щелей решетки, что ясно уже по самому способу его вывода. То же самое справедливо в отношении условия (30.42) для главных максимумов. Но здесь следует иметь в виду, что при определенных соотношениях между  $a$  и  $b$  некоторые главные максимумы могут и не появиться, если условие главных максимумов совпадет с условием главных минимумов. Так, при равенстве ширины  $a$  щели и расстояния  $b$  между щелями период решетки  $d = 2a$ . Условие главных максимумов принимает тогда вид  $a \sin \varphi = \pm m\lambda/2$ , что следует сравнить с условием главных минимумов  $a \sin \varphi = \pm m\lambda$ . Мы видим, что все максимумы четных порядков не могут возникнуть, так как соответствующие направления совпадают с направлениями, куда ни одна из щелей (а значит, и решетка в целом) лучей не посылает.

Максимальное значение синуса равно единице (достигается при максимальном угле дифракции  $\varphi = \pi/2$ ), поэтому максимальное значение порядка спектра равно  $m_{\max} = [d/\lambda]$ . Так, при  $d = 0,8$  мкм и  $\lambda = 500$  нм получаем оценку  $m_{\max} = [1,6] = 1$ .

Положение главных максимумов зависит от длины волны  $\lambda$ . Поэтому при пропускании через решетку белого света все максимумы, кроме центрального, разлагаются в спектр, фиолетовый конец которого обращен к центру дифракционной картины, а красный — наружу. Таким образом, дифракционная решетка представляет собой спектральный прибор, что и подметил первым Дж. Грегори. Заметим, что, в то время как спектральная призма сильнее всего отклоняет фиолетовые лучи, дифракционная решетка, наоборот, сильнее отклоняет красные лучи.

Как мы видели на примере дифракции на двух щелях, помимо главных минимумов, на дифракционной картине возникают и дополнительные минимумы. Их число между главными максимумами равно  $N - 1$ , а **условие дополнительных минимумов** имеет вид

$$\boxed{\begin{aligned} d \sin \varphi &= \pm \lambda \left( m + \frac{s}{N} \right), \\ s &= 1, 2, \dots, N - 1; \quad m = 0, 1, 2, \dots \end{aligned}} \quad (30.44)$$

Дополнительные минимумы разделены вторичными максимумами, создающими относительно слабый фон, их число между главными максимумами равно  $N - 2$ .

Это следует из распределения интенсивности для дифракционной решетки с  $N$  щелями, к выводу которого мы переходим. Амплитуда колебаний, приходящих на экран от одной щели, равна  $A_1$ . Колебания от соседней щели запаздывают по фазе (или опережают — это все равно) на  $k\Delta$  (см. рис. 30.8), от следующей — на  $2k\Delta$  и т. д. ( $k = 2\pi/\lambda$  — модуль волнового вектора). Складывая колебания от  $N$  щелей, находим:

$$A_N = A_1 \sum_{n=0}^{N-1} e^{-ink\Delta} = A_1 \frac{1 - e^{-iNk\Delta}}{1 - e^{-ik\Delta}}, \quad (30.45)$$

где применен известный результат для суммы геометрической прогрессии. Квадрат модуля амплитуды дает интенсивность колебаний  $I_N = |A_N|^2$  в заданном направлении:

$$I_N = I_1 \frac{(1 - e^{-iNk\Delta})(1 - e^{iNk\Delta})}{(1 - e^{-ik\Delta})(1 - e^{ik\Delta})} = I_1 \frac{1 - \cos(Nk\Delta)}{1 - \cos(k\Delta)} = I_1 \frac{\sin^2(Nk\Delta/2)}{\sin^2(k\Delta/2)}. \quad (30.46)$$

Интенсивность  $I_1$  излучения, исходящего от одной щели, выражается формулой (30.33), а разность хода  $\Delta$  лучей от соседних щелей — формулой (30.38). Вместе с (30.46) это приводит к окончательному выражению для распределения интенсивности света, распространяющегося под углом  $\varphi$ :

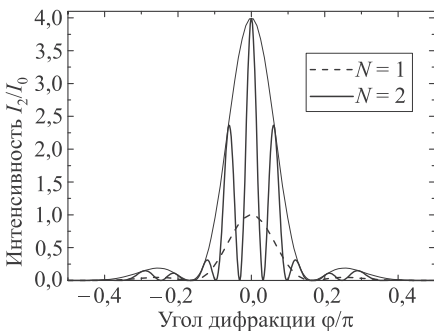
$$I_N = I_0 \frac{\sin^2 \psi}{\psi^2} \cdot \frac{\sin^2(N\delta)}{\sin^2 \delta}, \quad (30.47)$$

$$\psi = \pi \frac{a}{\lambda} \sin \varphi, \quad \delta = \pi \frac{d}{\lambda} \sin \varphi.$$

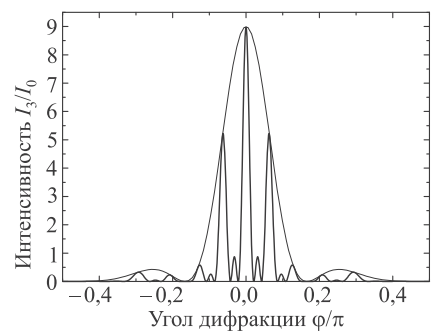
Интенсивность  $I_{N0}$  в центральном максимуме дифракционной решетки в  $N^2$  раз превышает интенсивность  $I_0$  в центральном максимуме от одной щели. Первый сомножитель в (30.47) дает распределение интенсивности от одной щели (ср. с (30.33)); он зависит от размера щелей, но не от параметров решетки в целом. Этот сомножитель называется *дифракционным членом*. Второй сомножитель зависит уже от периода решетки, но не от размера щелей или от промежутка между ними по отдельности. Он описывает интерференцию пучков света, приходящих от всех щелей решетки, и называется *интерференционным членом*.

Из формулы (30.47) сразу следуют условия дополнительных минимумов (30.44). Действительно, интерференционный член обращается в ноль, если  $\sin(N\delta) = 0$ , но  $\sin(\delta) \neq 0$ . Тогда  $N\delta$  должно быть кратно числу «пи» ( $N\delta = m'\pi$ ), но отношение  $m'/N$  не должно быть целым числом. Значит, при делении  $m'$  на  $N$  должен получаться остаток  $s$  — целое число от 1 до  $N - 1$ . Следовательно,  $m'$  можно представить в виде  $m' = Nm + s$ , откуда и вытекают условия (30.44).

Сказанное иллюстрируется графиками распределения интенсивности света при дифракции на двух (рис. 30.9) и трех (рис. 30.10) щелях.



**Рис. 30.9.** Распределение интенсивности при дифракции на двух щелях шириной  $a = 2\lambda$  при расстоянии между ними  $b = 3\lambda$ ; пунктиром показана дифракция на одной щели той же ширины; тонкая сплошная линия — увеличенная в 4 раза интенсивность для одной щели



**Рис. 30.10.** Распределение интенсивности при дифракции на трех щелях при тех же параметрах решетки, что и на рис. 30.9; здесь тонкая сплошная линия — распределение интенсивности для одной щели, увеличенное в 9 раз

**Задача 30.20.** В дифракционной решетке  $N$  щелей, каждая из которых короче предыдущей:  $l_{i+1} = \alpha l_i$ ,  $i = 1, \dots, N-1$ ;  $\alpha < 1$ . Какова интенсивность светового потока в центральном пике? Найдите выражение для распределения интенсивности в такой решетке при  $N \gg 1$ .

**Решение.** Уменьшение ширины щели ведет к пропорциональному уменьшению амплитуды исходящего из нее светового пучка. Поэтому выражение (30.45) будет теперь выглядеть так:

$$A_N = A_1 \sum_{n=0}^{N-1} \alpha^n e^{-ink\Delta} = A_1 \frac{1 - \alpha^N e^{-iNk\Delta}}{1 - \alpha e^{-ik\Delta}}. \quad (30.48)$$

Квадрат модуля этого выражения дает нам интенсивность светового потока:

$$I_N = I_1 \frac{1 - 2\alpha^N \cos(Nk\Delta) + \alpha^{2N}}{1 - 2\alpha \cos(k\Delta) + \alpha^2}. \quad (30.49)$$

В центральном пике  $\Delta = 0$  и  $I_1 = I_0$ , так что (30.49) дает

$$I_{N0} = I_0 \left( \frac{1 - \alpha^N}{1 - \alpha} \right)^2. \quad (30.50)$$

При большом числе щелей в решетке  $\alpha^N \rightarrow 0$  и мы вместо (30.49) получаем выражение

$$I_N = \frac{I_1}{1 - 2\alpha \cos(2\pi\Delta/\lambda) + \alpha^2} = \frac{I_1}{(1 - \alpha)^2 + 4\alpha \sin^2 \delta}. \quad (30.51) \quad \blacksquare$$

## Характеристики дифракционной решетки

Будучи оптическим прибором, дифракционная решетка характеризуется несколькими важными параметрами, численные значения которых дают возможность судить о ее применимости и эффективности в тех или иных технических и научных приложениях.

### Разрешающая способность дифракционной решетки

Важной характеристикой всякого спектрального прибора является *разрешающая способность*. Разрешающей способностью спектрального прибора называют безразмерную величину

$$R = \frac{\lambda}{\Delta\lambda}, \quad (30.52)$$

где  $\Delta\lambda$  — минимальная разность длин волн двух спектральных линий, при которой эти линии воспринимаются раздельно. Определим разрешающую способность дифракционной решетки. Положение середины максимума  $m$ -го порядка для длины волны  $\lambda + \Delta\lambda$  определяется условием

$$d \sin \varphi_{\max} = m(\lambda + \Delta\lambda). \quad (30.53)$$

Края этого максимума (т. е. ближайшие дополнительные минимумы) для длины волны  $\lambda$  расположены под углами, удовлетворяющими соотношению

$$d \sin \varphi_{\min} = \left( m \pm \frac{1}{N} \right) \lambda. \quad (30.54)$$

Два близких максимума воспринимаются раздельно в том случае, если середина одного максимума совпадает с краем другого (*критерий Рэлея*).

Таким образом, середина максимума для длины волны  $(\lambda + \Delta\lambda)$  совпадает с краем максимума для длины волны  $\lambda$  в том случае, если

$$m(\lambda + \Delta\lambda) = \left(m + \frac{1}{N}\right)\lambda. \quad (30.55)$$

Отсюда находим:

$$m\Delta\lambda = \frac{\lambda}{N}. \quad (30.56)$$

Следовательно, разрешающая способность дифракционной решетки

$$R = \lambda/\Delta\lambda = mN \quad (30.57)$$

пропорциональна порядку спектра  $m$  и числу щелей  $N$ .

**Задача 30.21.** Желтая линия в спектре излучения натрия на самом деле является дублетом — состоит из двух близких линий, соответствующих длинам волн  $\lambda_1 = 589, 592$  нм и  $\lambda_2 = 588, 995$  нм. Для изучения спектра используется дифракционная решетка, свет падает на нее по нормали. Каким должно быть минимальное число щелей решетки, чтобы с ее помощью можно было различить две линии дублета натрия?

**Решение.** Средняя длина волны в дублете равна  $\lambda = 589, 293$ , расщепление линий  $\Delta\lambda = \lambda_1 - \lambda_2 = 0, 597$  нм. Отсюда находим необходимую разрешающую способность дифракционной решетки:  $R = \lambda/\Delta\lambda = 987$ . В соответствии с формулой (30.57) получаем, что для разрешения линий дублета натрия в первом порядке ( $m = 1$ ) решетка должна содержать не менее  $N = 987$  штрихов. ■

### Угловая дисперсия дифракционной решетки

Еще одним важным параметром дифракционной решетки является ее *угловая дисперсия*, характеризующая угловое разрешение волн близкой частоты. Пусть для волны длиной  $\lambda$  угол дифракции, соответствующий какому-то из главных максимумов, равен  $\varphi$ , а для волны длиной  $\lambda + d\lambda$  он становится равным  $\varphi + d\varphi$ . Угловой дисперсией дифракционной решетки называется величина

$$D = \frac{d\varphi}{d\lambda}. \quad (30.58)$$

Дифференцируя условие главных максимумов (30.42) по  $\lambda$ , находим:

$$D = \frac{m}{d \cos \varphi} = \frac{1}{\sqrt{(d/m)^2 - \lambda^2}}. \quad (30.59)$$

Напомним, что здесь  $m$  — порядок максимума, а  $d$  — период решетки. Видно, что угловая дисперсия растет при увеличении порядка спектра и скользящих углов наблюдения, близких к  $\pi/2$ . Поэтому радужные блики видны даже на поверхности старинной грампластинки, если смотреть на источник света вдоль ее поверхности.



**Задача 30.22.** Угловая разрешающая способность человеческого глаза примерно равна одной угловой минуте:  $\Delta\varphi = 1'$ . В спектре какого порядка человек сможет раздельно увидеть близкие цвета — зеленый ( $\lambda_з = 560$  нм) и желтый ( $\lambda_ж = 580$  нм) — при дифракции белого света на дифракционной решетке с плотностью 100 штрихов на 1 см (примерная плотность звуковых канавок на грампластинках)? Каким должен быть при этом угол дифракции?

**Решение.** Переведем угловую разрешающую способность глаза в радианы:

$$\Delta\varphi = 1/60 \text{ град} = (\pi/180)/60 = 0,29 \cdot 10^{-3} \text{ рад.}$$

Разность длин волн в условии задачи  $\Delta\lambda = \lambda_ж - \lambda_з = 20$  нм  $= 2 \cdot 10^{-8}$  м. Отсюда следует, что угловая дисперсия дифракционной решетки должна превышать значение  $D = \Delta\varphi/\Delta\lambda = 14\,500 \text{ м}^{-1}$ .

Период нашей дифракционной решетки  $d = 1 \text{ см}/100 = 10^{-4}$  м. Из выражения (30.59) следует тогда нижняя граница для порядка спектра:

$$m \geq \frac{d}{\sqrt{1/D^2 + \lambda^2}} \approx dD = 1,45. \quad (30.60)$$

В выражении под квадратным корнем в знаменателе мы имели право пренебречь слагаемым  $\lambda^2$ , так как в нашей задаче  $1/D \gg \lambda$ .

Таким образом, зеленый и желтый цвета глаз может различить уже во втором порядке спектра:  $m = 2$ . Из (30.59) следует тогда неравенство

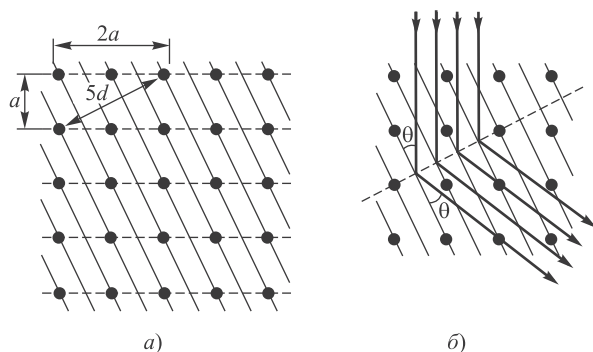
$$\cos\varphi \leq \frac{1}{dD} = 1/1,45 = 0,69, \quad (30.61)$$

т. е. угол дифракции должен превышать значение  $\varphi = \arccos 0,69 = 46,4^\circ$ . ■

## 30.6 Дифракция рентгеновских лучей

Вильгельм Конрад Рентген стал в 1901 г. первым лауреатом Нобелевской премии по физике «в знак признания исключительных услуг, которые он оказал науке открытием замечательных лучей, названных впоследствии в его честь». Сам Рентген предпочитал называть открытое им в 1895 г. излучение X-лучами («икс-лучами»), и это название укоренилось в ряде стран. Рентгеновские лучи — это электромагнитное излучение с длинами волн  $10^{-8} - 10^{-12}$  м, т. е. лежащее между ультрафиолетовым излучением и гамма-излучением. Дифракцию столь коротких волн невозможно наблюдать с помощью обычных дифракционных решеток.

К примеру, пустой DVD-диск имеет на поверхности спиралевидную дорожку для направления луча лазера, причем расстояние между соседними витками равно  $d = 0,74$  мкм  $= 740$  нм. Такой диск представляет собой отражательную дифракционную решетку с периодом  $d$ . Из условия главных максимумов (30.42) следует, что при использовании света с длиной волны  $\lambda = 550$  нм (желто-зеленый свет, центр видимого диапазона) угловое расстояние между центральным максимумом и максимумом первого порядка равно  $\varphi = \arcsin(\lambda/d) = \arcsin(550/740) = 48^\circ$ . Такой максимум легко наблюдается,



**Рис. 30.11.** В одном и том же кристалле с постоянной решетки  $a$  отражающие брэгговские плоскости можно провести по-разному; на рисунке *а*) показаны два варианта: семейство плоскостей с расстоянием между ними  $d = a/\sqrt{5}$  (сплошные линии) и  $d = a$  (штриховые линии); соответственно угол  $\theta$  скольжения луча будет разным для различных плоскостей: на рисунке *б*) показано отражение для наклонного расположения брэгговских плоскостей, для второго варианта  $\theta = 90^\circ$

и мы действительно видим игру цветов радуги, поворачивая плоскость диска. Но для рентгеновских лучей с длиной волны  $\lambda = 0,1$  нм в середине указанного выше диапазона угловое расстояние первого максимума составляет всего  $\varphi = \arcsin(\lambda/d) = \arcsin(0,1/740) = 0,0077^\circ$ . Это слишком малый угол для практического использования. Длины волн рентгеновского диапазона близки к размерам атомов, так что искусственным путем невозможно создать подходящие дифракционные решетки.

Однако природа позаботилась о нас и сама создала такие решетки. В кристаллах атомы расположены регулярно и образуют периодическую пространственную структуру — кристаллическую решетку. Межатомные расстояния в кристаллической решетке имеют тот же порядок, что и размеры атомов, и потому кристалл может служить дифракционной решеткой для рентгеновских лучей. Нобелевская премия за 1914 г. была присуждена немецкому ученому Макс фон Лауэ «за открытие дифракции рентгеновских лучей на кристаллах». Это открытие подтвердило волновую природу рентгеновского излучения и позволило определить структуру многих кристаллов.

Таковыми исследованиями занимались британские ученые Уильям Генри Брэгг и его сын Уильям Лоренс Брэгг, которые уже в следующем, 1915 г. получили Нобелевскую премию «за заслуги в исследовании структуры кристаллов с помощью рентгеновских лучей». Так называемое *брэгговское отражение* — специальный случай дифракции рентгеновских лучей. Кристаллическую решетку (например, кубическую решетку кристалла поваренной соли) можно рассматривать как семейство параллельных плоскостей, расположенных на расстоянии  $d$  друг от друга. В одной и той же решетке можно получить множество различных семейств (рис. 30.11). В каждом из них расстояние между плоскостями будет по-разному связано с периодом  $a$  кристаллической решетки, а сами плоскости будут по-разному наклонены к падающим на кристалл лучам.

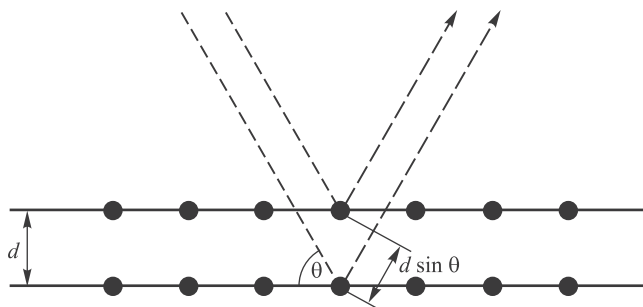


Рис. 30.12. К выводу условия Вульфа—Брэгга

Направление рассеяния рентгеновских лучей при отражении от семейства плоскостей, разделенных расстоянием  $d$ , определяется **законом Брэгга**<sup>2</sup>

$$2d \sin \theta = k\lambda, \quad k = 1, 2, 3, \dots \quad (30.62)$$

Вывод этого условия аналогичен выводу условий интерференции проходящих лучей в тонкой пленке и иллюстрируется на рис. 30.12. Отражение падающих лучей в обратном направлении происходит от брэгговских плоскостей, показанных на рис. 30.11 пунктирными линиями. Тогда угол скольжения  $\theta = 90^\circ$  и условие максимумов Брэгга принимает вид

$$2a = k\lambda, \quad k = 1, 2, 3, \dots \quad (30.63)$$

## 30.7 Голография

**Голография** — это запись и воспроизведение трехмерного изображения объекта. Как известно, световая волна характеризуется амплитудой и фазой. Обычная двумерная фотография основана на записи распределения интенсивности (квадрата амплитуды) световой волны, приходящей от фотографируемого объекта. Открытие способов записи и воспроизведения не только амплитуды, но и фазы световой волны, рассеянной на объекте, позволило получить его трехмерное изображение. Голографию изобрел в 1947 г. английский физик Д. Габор, он же назвал запись изображения объекта *голограммой* (от греч. *holos* — полный).

В 1971 г. Д. Габору была присуждена Нобелевская премия «за изобретение и усовершенствование голографического метода». В своей нобелевской лекции<sup>3</sup> он сказал, что находится в более выгодном положении, чем большинство его предшественников, потому что ему не нужно выписывать уравнения или показывать

<sup>2</sup>Эти условия были выведены У.Л. Брэггом (1912) и независимо советским кристаллографом Г.В. Вульфом (1913), поэтому в отечественной литературе закон (30.62) называется законом Вульфа—Брэгга.

<sup>3</sup>Д. Габор. Голография (1948–1971). Успехи физических наук, т. 100, вып. 1 (1973), с. 5.

сложные графики, так как сущность голографии можно объяснить и понять, рассматривая явления в рамках волновой теории света. Свой рассказ об истории открытия голографии в 1947 г. Габор начинает с двухщелевого опыта Т. Юнга, обнаружившего явление интерференции и внесшего существенный вклад в подтверждение и развитие знаменитого принципа Гюйгенса, согласно которому для воспроизведения распределения света в пространстве достаточно знать волновое поле на какой-то поверхности. В дальнейшем важные исследования в этой области были выполнены О. Френелем. Советский физик Ю. Н. Денисюк, один из основоположников оптической голографии, даже утверждал, что работы Юнга и Френеля заложили теоретические основы голографии и открытие этого явления могло быть сделано уже ими в самом начале XIX в.

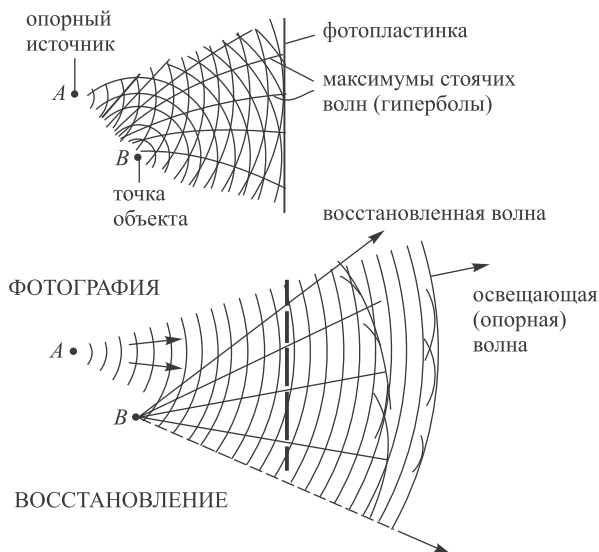
Сам Габор говорит, что в своих исследованиях он опирался на работы двух крупнейших физиков. Один из них — уже известный нам нобелевский лауреат У. Л. Брэгг, другой — голландский физик Ф. Цернике, лауреат нобелевской премии (1953) «за обоснование фазово-контрастного метода, особенно за изобретение фазово-контрастного микроскопа». О предыстории голографии можно прочитать в статье<sup>4</sup>.

Трудно изложить основную идею голографического метода прозрачнее, чем это сделал отец голографии Д. Габор в своей нобелевской лекции, взяв в качестве примера простейший объект — точку (рис. 30.13).

«Как известно, в обычной фотографии фаза полностью теряется, а регистрируется только интенсивность. Потеря фазы не вызывает удивления, поскольку в этом случае не существует ничего, с чем можно было бы ее сравнить! Давайте посмотрим, что произойдет, если ввести эталон для фазы, добавить так называемый «когерентный фон». Моя аргументация будет понятна на простейшем примере, представленном на рисунке для точечного объекта. Благодаря интерференции предметной волны с когерентным фоном, или «опорной волной», возникнут интерференционные полосы. Всюду, где фазы двух волн оказываются одинаковыми, будут образовываться максимумы, соответствующие повышению яркости. Если изготовить фототранспарант этой картины таким образом, чтобы свет проходил через него в тех местах, где находятся максимумы интерференционной картины, и затем осветить его только одним опорным пучком, то, безусловно, фазы поля будут соответствовать фазам опорного источника  $A$ . Однако, поскольку в прозрачных участках фотопозитива фазы поля источника  $A$  и фазы поля от объекта  $B$  были одинаковыми, в этих пропускающих участках фазы поля будут одновременно соответствовать фазам поля от объекта  $B$ . Это означает, что с неизбежностью появится *восстановленная* волна  $B$ ».

Рассмотренный принцип работает при формировании изображений от любой системы точек, т. е. от объектов любой сложности. Голограмма содержит не одно

<sup>4</sup>Ю. Н. Денисюк, В. А. Гуриков. Развитие голографии как нового научного направления. Природа, 1984, № 5, с. 54.



**Рис. 30.13.** Основная идея голографии (рисунок из нобелевской лекции Д. Габора)

изображение объекта, а множество в зависимости от того, под каким углом она рассматривается. Первые голограммы были получены при расположении всех элементов на одной оси. Голограмма формировала два изображения — мнимое в точке нахождения объекта и реальное в оптически сопряженной точке, и оба находились в опорном пучке света. Вследствие малой когерентности источников света в те годы можно было реализовать лишь такую схему, страдающую неустранимыми искажениями получающихся изображений. Изобретались изощренные методы избавления от второго изображения, но примерно к 1955 г. голография впала в спячку, по образному выражению Габора.

Из такого состояния эту область (тогда науки, теперь еще и техники) вывело открытие лазеров. Высокая когерентность лазерного излучения привела к созданию схемы с наклонным опорным пучком, что позволило разнести по углам и глубине мнимое и действительное изображения. К тому же интенсивность лазерного излучения намного превышала интенсивность света ртутной лампы, использовавшейся ранее, что позволило получать голограммы больших размеров с достаточно высоким разрешением. Первая лазерная голограмма была получена в 1963 г., с этих пор развитие голографии приобрело взрывной характер.

В 1962 г., накануне этого «взрыва», Ю. Н. Денисюк опубликовал важную работу. Он обратил внимание на то, что двумерная голограмма представляет собой частный случай более общего явления. Интерференционная картина возникает не только на регистрирующем экране (фотопластинке), но и во всем трехмерном пространстве в виде стоячих волн. Если их «сфотографировать», то получившаяся голограмма воспроизведет не только форму объекта, но и его цвета. Денисюк отмечал, что его идея базировалась на работах Г. Липпмана,

получившего Нобелевскую премию по физике в 1908 г. «за создание метода фотографического воспроизведения цветов на основе явления интерференции». По способу Липпмана в фотоматериал впечатывалась картина стоячих волн, отраженных от плоского зеркала. Подобно этому, по методу Денисюка перед объектом со стороны источника света устанавливалась фотопластинка с толстым эмульсионным слоем. Часть опорной волны проходит через фотопластинку и освещает объект. Рассеянный от объекта свет падает на пластинку с другой стороны от опорного луча и интерферирует с ним. После экспозиции и проявления в фотопластинке образуется трехмерная слоистая структура, моделирующая пространственное распределение интенсивности в стоячей волне, образованной при наложении излучения источника и света, рассеянного от объекта. Объемная модель стоячих волн (голограмма) однозначно воспроизводит амплитуду, фазу и спектральный состав записанного излучения. Если осветить голограмму излучением любого источника со сплошным («белым») спектром (солнечным светом или даже обычной лампой накаливания), то возникнет цветное пространственное изображение объекта.

Как писал в 1962 г. Ю. Н. Денисюк, «данное явление может оказаться полезным для развития изобразительной техники, создающей полную иллюзию действительности изображаемых объектов, в структурном анализе, гидролокации, радиолокации, ультразвуковой дефектоскопии, а также для изготовления диспергирующих элементов типа дифракционной решетки».

## Контрольные вопросы

1. Обсудите сходство и различие между интерференцией и дифракцией.
2. Почему частицы размером 0,3 мкм неразличимы в микроскоп?
3. Мы слышим звуки, доносящиеся из-за угла, но не видим их источника. Объясните это различие в распространении звуковых и световых волн.
4. Объясните метод зон Френеля.
5. При каких условиях в центре тени от круглого диска наблюдается светлое пятно? А при каких условиях в центре дифракционной картины от круглого отверстия наблюдается темное пятно?
6. Сформулируйте физические соображения, приводящие к условиям дифракционных максимумов и минимумов при дифракции Фраунгофера на щели.
7. Как меняется ширина центрального дифракционного максимума при изменении ширины щели?
8. Опишите различия дифракционных картин при дифракции на одной и двух щелях.
9. Какие физические явления происходят при падении светового потока на многощелевую дифракционную решетку?
10. Почему дифракционная решетка является спектральным прибором?

11. При изготовлении перламутровых пуговиц на их поверхность наносится мельчайшая штриховка. Почему после такой обработки пуговица имеет радужную окраску?
12. Что такое разрешающая способность дифракционной решетки?
13. Что такое угловая дисперсия дифракционной решетки?
14. Опишите явление брэгговского отражения рентгеновских лучей.
15. Какова основная идея создания и воспроизведения трехмерных голографических изображений?

## Глава 31

# Влияние среды на свойства света

В этой главе мы рассмотрим некоторые эффекты, возникающие при распространении света в среде. Первый из них — *поляризация* — означает поперечную анизотропию световых волн, т. е. различие свойств света в разных направлениях, перпендикулярных световому лучу. Далее мы рассмотрим *дисперсию* световых волн в среде, проявляющуюся в зависимости скорости света в среде от его частоты (длины волны). Наконец, мы коротко остановимся на *поглощении* и *рассеянии* света в среде.

### 31.1 Поляризация света

Следствием теории Максвелла является поперечность электромагнитных (световых) волн: векторы напряженности электрического и магнитного полей волны взаимно перпендикулярны и колеблются перпендикулярно вектору скорости  $\vec{v}$  распространения волны (т. е. перпендикулярно световому лучу). Явление поляризации света служит надежным обоснованием поперечности световой волны. При рассмотрении поляризации обычно все рассуждения связаны с плоскостью колебаний вектора напряженности электрического поля  $\vec{E}$  (его называют световым вектором), так как химическое, физиологическое и другие виды воздействия света на вещество обусловлены главным образом электрическими колебаниями. Однако при этом следует помнить об обязательном существовании перпендикулярного ему вектора напряженности магнитного поля  $\vec{H}$ .

#### Поляризация электромагнитной волны

Со свойствами электромагнитных волн мы познакомимся ближе в следующей главе. Сейчас же нам достаточно того, что электрическое поле плоской электромагнитной волны записывается как

$$\vec{E}(\vec{r}, t) = \vec{E}_0 \cos(\omega t - \vec{k} \cdot \vec{r} + \varphi_0). \quad (31.1)$$

Если направление вектора  $\vec{E}_0$ , характеризующего амплитуду колебаний электрического поля, не зависит от времени, то вектор электрического поля  $\vec{E}(\vec{r}, t)$  всегда



и во всех точках колеблется вдоль одной прямой линии. Плоскость, содержащая эту прямую и направление распространения волны, называется **плоскостью поляризации**, а волна — **линейно-поляризованной** (или, что то же самое, **плоскополяризованной**).

Выбирая ось  $x$  вдоль направления распространения волны, а ось  $y$  вдоль векторной амплитуды  $\vec{E}_0$ , записываем (31.1) в виде:

$$\vec{E}_y(\vec{r}, t) = \vec{n}_y E_{0,y} \cos(\omega t - kx + \varphi_{0,y}). \quad (31.2)$$

Однако существует и вторая линейно-поляризованная волна, имеющая ту же частоту и распространяющаяся в том же направлении:

$$\vec{E}_z(\vec{r}, t) = \vec{n}_z E_{0,z} \cos(\omega t - kx + \varphi_{0,z}). \quad (31.3)$$

Электрические колебания в этой волне направлены вдоль оси  $z$ , так что волны (см. (31.2) и (31.3)) линейно независимы. Обе они являются решением одного и того же волнового уравнения, так что их суперпозиция также является решением того же уравнения. Сложив эти волны, мы найдем общее выражение для монохроматической волны с данной частотой  $\omega$ , распространяющейся вдоль оси  $x$ . Математически эта процедура ничем не отличается от сложения взаимно ортогональных колебаний. Если зафиксировать какую-то точку  $x$  и следить за изменением вектора электрического поля в ней, то конец вектора будет описывать в общем случае эллиптическую траекторию в плоскости, параллельной  $yOz$ . Вращение  $\vec{E}$  происходит с частотой волны  $\omega$ . При равенстве амплитуд  $E_{0,y}$  и  $E_{0,z}$  эллипс превращается в окружность. Тогда говорят о **круговой поляризации** волны. В соответствии с двумя возможными направлениями вращения вектора  $\vec{E}(\vec{r}, t)$  возможны право- и левополяризованные волны. Любую другую волну можно представить как линейную комбинацию двух линейно-поляризованных волн или как линейную комбинацию двух волн с круговой поляризацией. Иными словами, электромагнитные волны имеют две внутренние степени свободы.

## Естественный и поляризованный свет

Свет от естественного источника является суперпозицией линейно-поляризованных волн, испущенных большим числом независимых атомарных излучателей. В суммарной плоской волне световые векторы отдельных составляющих ортогональны направлению распространения волны (выбранному, например, в качестве оси  $x$ ), но по-разному ориентированы в плоскости  $yOz$ , причем все эти ориентации равновероятны. Такой свет называется **естественным** или неполяризованным.

Если под влиянием внешних воздействий на свет или внутренних особенностей источника света (лазер) появляется предпочтительное, наиболее вероятное направление колебаний, то такой свет называется **частично поляризованным**. неполяризованный (естественный) свет может испускаться лишь огромным числом элементарных излучателей. Электромагнитная волна от отдельного элементарного излучателя (атома, молекулы) всегда поляризована. С помощью спектральных устройств из пучка естественного света можно выделить волны, в которых колебания вектора  $\vec{E}$  будут происходить в одном

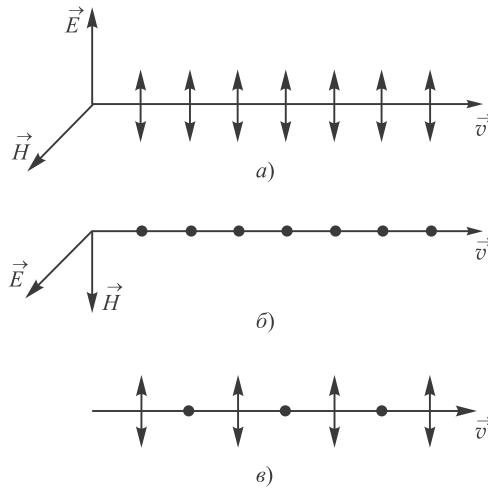


Рис. 31.1. Условные обозначения для поляризации волны

определенном направлении, перпендикулярном лучу, т. е. суммарная волна будет линейно-поляризованной.

Направление колебаний электрического поля линейно-поляризованной волны изображается следующим образом. Если вектор  $\vec{E}$  колеблется в плоскости чертежа, то на направление вектора скорости  $\vec{v}$  наносится ряд вертикальных черточек (рис. 31.1, а), а если в плоскости, перпендикулярной чертежу, — ряд точек (рис. 31.1, б). Естественный луч условно обозначается чередующимися черточками, которым соответствует, например, компонента  $E_y$  вектора напряженности электрического поля, и точками, соответствующими другой компоненте  $E_z$  (рис. 31.1, в).

Существуют приборы (поляризаторы), пропускающие только колебания, происходящие параллельно некоторой плоскости, называемой **плоскостью поляризации** прибора, и полностью задерживающие ортогональные колебания. Если пропустить через такой прибор пучок света, то на выходе он будет линейно-поляризованным. При вращении прибора вокруг направления луча интенсивность выходящего света будет изменяться от  $I_{\max}$  до  $I_{\min}$ . **Степенью поляризации** света называют величину

$$P = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}. \quad (31.4)$$

Если падающий пучок света линейно-поляризован, то при положении прибора, когда его плоскость поляризации ортогональна плоскости колебаний волны, свет через прибор не пройдет, т. е.  $I_{\min} = 0$ . В соответствии с уравнением (31.4), степень поляризации такого света  $P = 1$ . Для частично поляризованного света  $I_{\min} \neq 0$  и  $P < 1$ . Для естественного света, где волны разных поляризаций смешаны в равной степени и все направления эквивалентны, интенсивность выходящего света не изменяется при вращении поляризатора, так что  $I_{\max} = I_{\min}$

и  $P = 0$ . Такой же результат получится для света, поляризованного по кругу. Тут важно подчеркнуть, что степень поляризации света введена для характеристики хаотичности в направлении колебаний различных смешанных в нем волн. При круговой же поляризации хаотичности нет, колебания происходят упорядоченно. Поэтому понятие степени поляризации к таким волнам просто не применяется.

## Закон Малюса

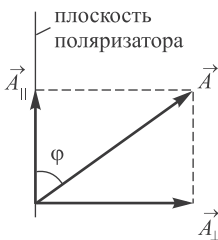
В качестве поляризаторов могут быть использованы среды, анизотропные в отношении колебаний вектора  $\vec{E}$ , например природные кристаллы турмалина. Монокристалл турмалина поглощает колебания вектора  $\vec{E}$  одного направления настолько сильно, что сквозь пластинку толщиной порядка 1 мм проходит только линейно-поляризованный луч. Кристаллы иодистого хинина еще сильнее поглощают одну из поляризаций: кристаллическая пленка толщиной в десятую долю миллиметра практически полностью отделяет один из линейно-поляризованных лучей.

Пусть естественный свет распространяется перпендикулярно плоскости рисунка 31.2. Амплитуду колебаний  $\vec{A}$ , совершающихся в плоскости, образующей с плоскостью поляризатора угол  $\varphi$ , можно разложить на два колебания с амплитудами  $A_{\parallel} = A \cos \varphi$  и  $A_{\perp} = A \sin \varphi$ .

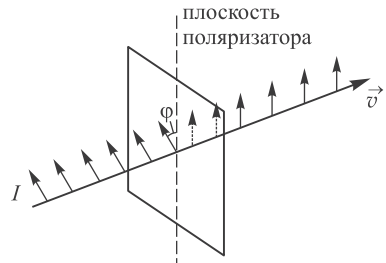
Первое колебание с амплитудой  $A_{\parallel}$  пройдет через прибор (поляризатор), второе — с амплитудой  $A_{\perp}$  — будет задержано (поглощено). Интенсивность прошедшей волны пропорциональна квадрату амплитуды  $A_{\parallel}^2 = A^2 \cos^2 \varphi$ . Падающая волна является смесью волн с различными  $\varphi$ . Усредняя по углам, получаем для интенсивности света на выходе из поляризатора

$$I_{\text{прош}} = I_0 \langle \cos^2 \varphi \rangle, \quad (31.5)$$

где  $I_0$  — интенсивность падающего на поляризатор света. В естественном свете все значения угла  $\varphi$  равновероятны:  $\langle \cos^2 \varphi \rangle = 1/2$ , так что интенсивность света, прошедшего через поляризатор, будет равна  $I_0/2$ . При вращении поляризатора вокруг направления луча естественного света интенсивность прошедшего света остается неизменной, но изменяется лишь ориентация плоскости колебаний света, выходящего из прибора.



**Рис. 31.2.** Разложение вектора амплитуды колебаний  $\vec{A}$  в волне, падающей на поляризатор



**Рис. 31.3.** Прохождение плоскополяризованной волны через поляризатор

Рассмотрим теперь падение линейно-поляризованного света с интенсивностью  $I_0$  на тот же поляризатор (рис. 31.3). Сквозь прибор пройдет составляющая колебаний с амплитудой  $A_{\parallel} = A_0 \cos \varphi$ , где  $\varphi$  — угол между плоскостью колебаний вектора  $\vec{E}$  и плоскостью поляризатора. Следовательно, интенсивность прошедшего света  $I$  определяется выражением

$$I = I_0 \cos^2 \varphi, \quad (31.6)$$

которое носит название *закона Малюса*.

Поляризационные приборы по своему целевому назначению делятся на *поляризаторы* и *анализаторы*. Поляризаторы служат для получения поляризованного света. С помощью анализатора можно убедиться, что падающий свет поляризован, и выяснить направление поляризации. Принципиальных конструктивных различий между поляризатором и анализатором не существует.

Поставим на пути естественного света два поляризатора, плоскости которых образуют угол  $\varphi$ . Из первого поляризатора выйдет линейно-поляризованный свет, интенсивность которого  $I_1$  составит половину интенсивности падающего естественного света  $I_0$ . Согласно закону Малюса из второго поляризатора (который играет роль анализатора) выйдет свет с интенсивностью  $I_2 = I_1 \cos^2 \varphi$ . Таким образом, интенсивность света, прошедшего через два поляризатора, равна:

$$I_2 = \frac{I_0}{2} \cos^2 \varphi. \quad (31.7)$$

Если угол  $\varphi = 0$  (плоскости поляризации поляризатора и анализатора параллельны), то  $I_2 = I_0/2$ ; если  $\varphi = \pi/2$  (анализатор и поляризатор скрещены), то  $I_2 = 0$ .

**Задача 31.23.** В частично поляризованном свете амплитуда колебаний, соответствующая максимальной интенсивности света при прохождении через поляризатор, в  $n = 2$  раза больше амплитуды, соответствующей минимальной интенсивности. Чему равна степень поляризации света?

**Решение.** Поскольку интенсивности пропорциональны квадратам амплитуд, имеем:  $I_{\max} = n^2 I_{\min}$ . Отсюда степень поляризации света равна:

$$P = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{n^2 - 1}{n^2 + 1} = 0,6. \quad \blacksquare$$

**Задача 31.24.** На пути света со степенью поляризации  $P = 0,6$  поставили анализатор так, что интенсивность прошедшего света стала максимальной. Во сколько раз уменьшится интенсивность, если анализатор повернуть на угол  $\alpha = 30^\circ$ ?

**Решение.** В падающем луче по условию (см. предыдущую задачу)  $I_{\max} = 4I_{\min}$ . При повороте анализатора на угол  $\alpha$  будут пропущены колебания, параллельные плоскости поляризации прибора. Поэтому интенсивность прошедших колебаний, прежде бывших параллельными плоскости поляризации, составит  $I_1 = I_{\max} \cos^2 \alpha$ , а интенсивность прошедших колебаний, до поворота задерживавшихся анализатором, равна  $I_2 = I_{\min} \sin^2 \varphi = (I_{\max}/4) \sin^2 \alpha$ . Суммарная интенсивность прошедших колебаний

равна сумме

$$\begin{aligned}
 I &= I_1 + I_2 = I_{\max} \left( \cos^2 \alpha + \frac{\sin^2 \alpha}{4} \right) = I_{\max} \left( 1 - \frac{3}{4} \sin^2 \alpha \right) = \\
 &= I_{\max} \left( 1 - \frac{3}{4} \cdot \frac{1}{4} \right) = \frac{13}{16} I_{\max}.
 \end{aligned} \tag{31.8}$$

Ответ: интенсивность уменьшится при повороте анализатора в  $16/13 = 1,23$  раза. ■

## Поляризация при отражении и преломлении

Получить поляризованный свет из естественного можно еще одним способом — отражением. Опыт показывает, что отраженный и преломленный лучи всегда частично поляризованы. Когда свет падает на неметаллическую поверхность, то в отраженном луче преобладают колебания, перпендикулярные плоскости падения (точки на рис. 31.4), а в преломленном луче — колебания, параллельные плоскости падения (стрелки на рис. 31.4). Степень поляризации зависит от угла падения лучей и коэффициента преломления отражающей среды. Исследуя это явление, британский физик Д. Брюстер установил, что при определенном значении угла падения  $i_{\text{п}} = i_{\text{Бр}}$ , удовлетворяющем условию

$$\boxed{\operatorname{tg} i_{\text{Бр}} = n_{21} = \frac{n_2}{n_1}}, \tag{31.9}$$

отраженный свет полностью поляризован в плоскости, перпендикулярной плоскости падения луча (плоскостью падения называется плоскость, содержащая падающий луч и нормаль к поверхности раздела двух сред). Это соотношение известно как **закон Брюстера**. При  $i_{\text{п}} = i_{\text{Бр}}$  отражается только та компонента вектора напряженности электрического поля, которая параллельна поверхности

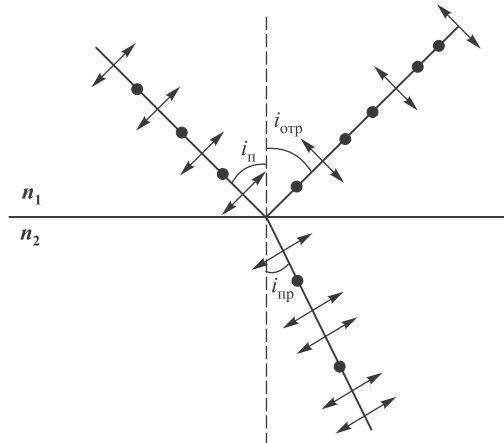
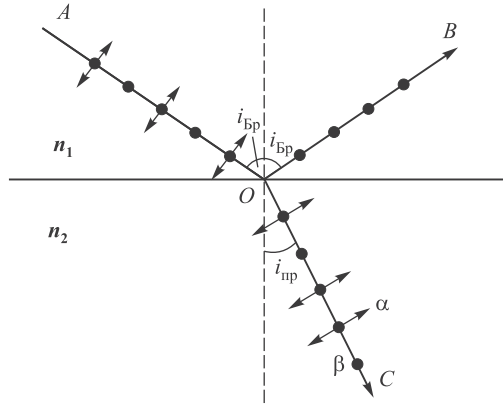


Рис. 31.4. Поляризация света при отражении и преломлении



**Рис. 31.5.** Ход лучей при падении света под углом Брюстера: отраженный луч ортогонален преломленному, поэтому излучатели типа  $\alpha$  не вносят вклад в поляризацию отраженного луча

диэлектрика (перпендикулярна плоскости падения). Соответственно преломленный луч всегда частично поляризован, так как отражается лишь какая-то доля падающего света.

*При падении света под углом Брюстера отраженный и преломленный лучи взаимно перпендикулярны:*  $i_{\text{пр}} = \pi/2 - i_{\text{Бр}}$ . Действительно, при  $i_{\text{п}} = i_{\text{Бр}}$  находим с учетом закона преломления:

$$\frac{\sin i_{\text{Бр}}}{\sin i_{\text{пр}}} = n_{21} = \frac{\sin i_{\text{Бр}}}{\cos i_{\text{Бр}}}. \quad (31.10)$$

Получаем отсюда

$$\cos i_{\text{Бр}} = \sin i_{\text{пр}}. \quad (31.11)$$

Таким образом,  $i_{\text{Бр}} + i_{\text{пр}} = \pi/2$ , откуда следует, что преломленный луч  $OC$  перпендикулярен отраженному лучу  $OB$  (рис. 31.5). Следует отметить, что на опыте закон Брюстера не выполняется вполне строго и неприменим в случае отражения света от поверхности проводников (металлов).

Физический смысл закона Брюстера можно объяснить следующим образом. Отраженный свет есть результат излучения вторичных волн электрическими зарядами (электронами) второй среды, колеблющимися под действием светового вектора волны. Эти колебания происходят в направлении светового вектора — электрического поля волны  $\vec{E}$ . Поскольку при падении луча на границу раздела двух сред под углом Брюстера отраженный луч, как было доказано, перпендикулярен преломленному, он не может содержать колебания светового вектора  $\vec{E}$ , происходящие в плоскости падения. Действительно, как видно из рис. 31.5, векторы колебаний типа  $\alpha$ , которые происходят в плоскости падения, параллельны отраженному лучу  $OB$ . Согласно электромагнитной теории Максвелла колеблющийся электрический заряд не излучает электромагнитных волн вдоль

направления своего движения (см. в разд. 32.10). Поэтому колеблющийся в диэлектрике излучатель типа  $\alpha$  не излучает вдоль направления  $OB$ . Таким образом, по направлению отраженного луча  $OB$  распространяется свет, посылаемый только излучателем типа  $\beta$ , направление колебаний которого перпендикулярно плоскости падения.

**Задача 31.25.** На какой угловой высоте  $\varphi$  над горизонтом должно находиться Солнце, чтобы свет, отраженный от поверхности воды, был полностью поляризован?

**Решение.** Угол падения света связан с угловой высотой Солнца над горизонтом соотношением  $i_{\text{п}} = \pi/2 - \varphi$ . По условию задачи угол падения равен углу Брюстера, так что  $\text{tg } i_{\text{п}} = \text{ctg } \varphi = n$ . Показатель преломления воды  $n = 1,33$ . Отсюда находим:  $\varphi = 37^\circ$ . ■

**Задача 31.26.** Угол Брюстера при падении света из воздуха на кристалл каменной соли равен  $i_{\text{Бр}} = 57^\circ$ . Определить скорость света  $v$  в этом кристалле.

**Решение.** Поскольку показатель преломления воздуха равен единице, показатель преломления каменной соли  $n$  совпадает с относительным показателем преломления  $n_{21}$  этих двух сред. Поэтому имеем:  $v = c/n = c/\text{tg}57^\circ = 0,65c = 1,95 \cdot 10^8$  м/с. ■

## 31.2 Дисперсия света

◆ **Дисперсией света** называется зависимость показателя преломления  $n$  вещества от длины волны света в вакууме  $\lambda$ :

$$n = n(\omega) = n(\lambda). \quad \blacklozenge \quad (31.12)$$

То же самое можно выразить по-другому: дисперсия света связана с зависимостью фазовой скорости световых волн от частоты:

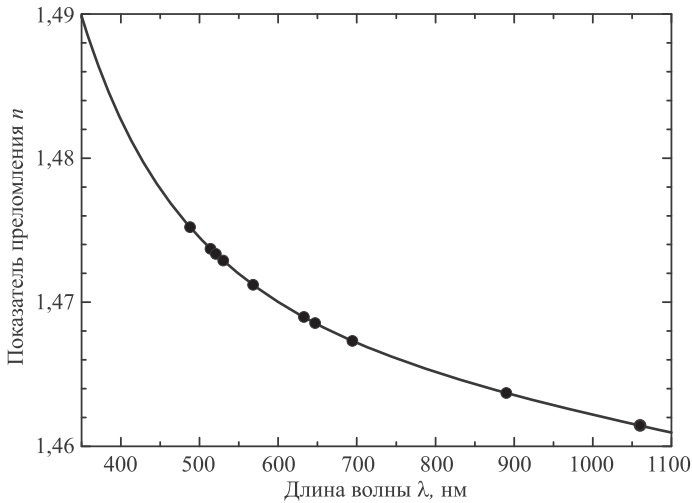
$$v = \frac{c}{n(\omega)}. \quad (31.13)$$

Для всех прозрачных бесцветных веществ в видимой части спектра с уменьшением длины волны (ростом частоты) показатель преломления увеличивается:  $dn/d\lambda < 0$  (или  $dn/d\omega > 0$ ). Такая дисперсия называется *нормальной*.

На рисунке 31.6 представлена кривая, описывающая экспериментальные данные для показателя преломления одной из марок оптического (кронового) стекла.

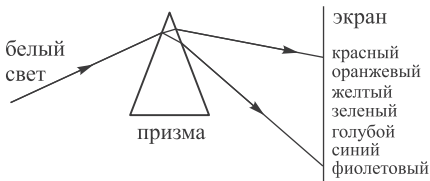
Изучением *нормальной дисперсии* занимался И. Ньютон. Разложение белого света в спектр при прохождении сквозь призму является следствием дисперсии света. При прохождении пучка белого света через стеклянную призму на экране возникает *разноцветный спектр* (рис. 31.7). Наибольшую длину волны и наименьший показатель преломления имеет красный свет, поэтому красные лучи отклоняются призмой слабее других. Большее отклонение будут испытывать последовательно лучи оранжевого, потом желтого, зеленого, голубого, синего и, наконец, фиолетового света. Происходит разложение падающего на призму сложного белого света на монохроматические составляющие.

Ярким проявлением дисперсии является радуга (впрочем, в этом природном явлении проявляются, как мы увидим, также и эффекты интерференции, дифракции и поляризации света). Радуга появляется при дожде и солнце и наблюдается в противоположной от солнца стороне (рис. 31.8). Кроме того, радуга

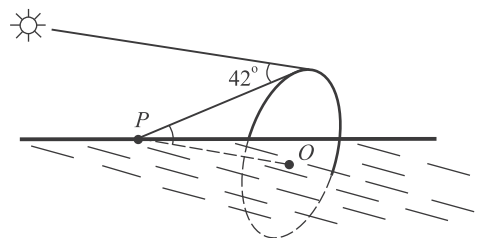


**Рис. 31.6.** Нормальная дисперсия оптического стекла (легкого крона ЛК6): точки — экспериментальные данные из справочника «Физические величины». М.: Энергоатомиздат, 1991. С. 770–771, кривая — интерполяционная формула, полученная на основе этих данных

видна в водопадах и может быть получена искусственно при разбрызгивании воды. Поэтому ученые давно поняли связь радуги с ходом солнечных лучей в дождевых каплях. Лучи от радуги составляют с первоначальным направлением солнечных лучей углы в интервале от  $40^\circ$  (для внутренней, фиолетовой дуги) до  $42^\circ$  (для внешней, красной дуги, рис. 31.7). На горизонтальной поверхности часть радуги скрыта от наблюдателя под горизонтом (пассажиры авиалайнеров иногда могут наблюдать полную окружность радуги с тенью от самолета в центре). При высоте солнца над горизонтом больше  $42^\circ$  радуга становится невидимой.



**Рис. 31.7.** При прохождении белого света через призму вследствие различия значений показателя преломления стекла для разных длин волн пучок разлагается на монохроматические составляющие — на экране возникает спектр



**Рис. 31.8.** Расположение радуги; точка  $P$  — местонахождение наблюдателя



Порядок цветов в радуге запоминается по известному мнемоническому правилу «каждый охотник желает знать, где сидит фазан», где первые буквы обозначают соответствующие цвета: красный, оранжевый, желтый, зеленый, голубой, синий, фиолетовый. На самом деле выделение семи цветов — чистая условность, потому что спектр радуги непрерывен. Да и сами основные цвета и их число различны у разных народов. Помимо яркой радуги, вся область неба, охватываемая ее дугой, выглядит ярче основного фона.

Нередко в небе наблюдаются две радуги, причем радуга второго порядка лежит вне первой — в интервале углов  $50^\circ$ – $53^\circ$ . Яркость ее гораздо меньше, а порядок цветов обратен порядку цветов в радуге первого порядка: внешняя дуга окрашена в фиолетовый цвет, а нижняя — в красный. Темный промежуток в  $8^\circ$  между двумя красными полосами в радугах первого и второго порядка называется *александровой полосой* по имени древнегреческого философа II в. н. э. Александра Афродисийского, впервые описавшего это явление. Александрова полоса гораздо темнее радуг, ее яркость соответствует яркости основного фона неба.

Геометрическую теорию радуги, объясняющую ее возникновение и расположение на небосводе, дал Марк Антоний де Доминис, хорватский ученый и архиепископ Сплита<sup>1</sup>. И. Ньютон называл его «знаменитым» и писал, что, согласно де Доминису, «внутренняя дуга образуется в круглых каплях дождя посредством двух преломлений солнечного света и одного отражения между ними, внешняя же дуга образуется посредством двух преломлений и двух отражений между ними в каждой капле воды; он доказывает свои объяснения при помощи опытов, произведенных с фиалом, наполненным водою, и со стеклянными шарами, наполненными водою и помещенными на солнце так, что были видны цвета обеих дуг»<sup>2</sup>. Таким образом, де Доминис вместе со своим современником Г. Галилеем может считаться одним из основателей экспериментальной физики. Р. Декарт в главе «О радуге» трактата «Метеоры»<sup>3</sup> (1637) описывает свои такие же опыты и приходит к аналогичному результату. Наконец, научное объяснение цветов в радуге дал И. Ньютон на основе своих экспериментов с призмами (опубликовано в 1672 г.). Суть объяснения в том, что и красные, и фиолетовые лучи преломляются сферическими капельками воды и отражаются от их внутренней поверхности. Но красные лучи преломляются меньше и попадают в глаз наблюдателя от капелек, находящихся на большей высоте. Поэтому верхняя полоса выглядит красной, а нижняя — фиолетовой. Известен афоризм, что Декарт поместил радугу в нужное место на небе, а Ньютон раскрасил ее.

**Задача 31.27.** На каких широтах  $\phi$  можно увидеть радугу в полдень в день летнего солнцестояния? Наклон земного экватора к эклиптике  $\epsilon = 23^\circ$ .

**Решение.** Схема расположения Земли показана на рис. 31.9. Пусть  $\phi$  — широта некоторой точки  $M$  на земной поверхности. Точка  $O$  — центр Земли,  $N$  и  $S$  — ее

<sup>1</sup>Жизнь де Доминиса завершилась в тюрьме инквизиции, где он и умер в 1624 г., не дождавшись приговора по обвинению в ереси, который был оглашен тремя месяцами позже. Тело де Доминиса было извлечено из гроба и публично сожжено на римской площади вместе с его книгами.

<sup>2</sup>И. Ньютон. Оптика. М.: Гостехиздат, 1954. С. 127.

<sup>3</sup>Р. Декарт. Рассуждения о методе. М.: АН СССР, 1953. С. 264.

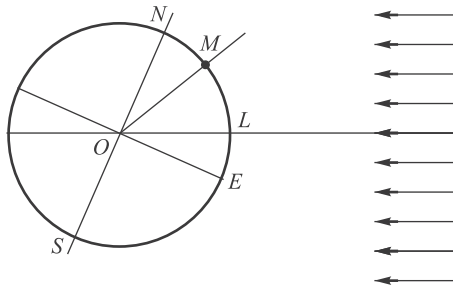


Рис. 31.9. К задаче 31.27

Северный и Южный полюсы,  $E$  — точка экватора, обращенная к Солнцу в указанное в условии время. Угол  $\angle EOL = \epsilon$  — наклон экватора к эклиптике, угол  $\angle EOM = \phi$  — широта точки  $M$ . Отсюда угол  $\angle LOM$  между направлением на Солнце (отрезок  $OL$ ) и вертикалью к поверхности в точке  $M$  (отрезок  $OM$ ) равен  $\angle LOM = \angle EOM - \angle EOL = \phi - \epsilon$ . Тогда высоту  $\theta$  Солнца над горизонтом можно записать как  $\theta = 90^\circ - \angle LOM = 90^\circ - \phi + \epsilon = 113^\circ - \phi$ . Радугу в точке  $M$  можно увидеть, если Солнце стоит не слишком высоко над горизонтом:  $\theta = 113^\circ - \phi \leq 42^\circ$ , откуда получаем  $\phi \geq 71^\circ$ . Это на пять градусов севернее Северного полярного круга (подходит, например, Шпицберген). ■

## Теория радуги

Итак, солнечный луч, попавший на дождевую каплю, может отразиться от нее или войти внутрь и выйти через заднюю стенку. Но нас интересует третий путь, когда луч преломляется внутри капли, отражается от ее задней стенки и, снова преломившись, выходит наружу через переднюю стенку (рис. 31.10). Самый простой способ вычисления угла поворота луча дан в статье<sup>4</sup>. При первом преломлении луч поворачивается на угол  $i - r$  (где  $i$  — угол падения луча на поверхность капли, а  $r$  — угол отражения). При отражении от задней стенки капли луч поворачивается на угол  $\pi - 2r$ . Наконец, выходя наружу, луч снова поворачивается на угол  $i - r$ . Таким образом, для полного угла поворота луча имеем

$$\theta_0 = \pi + 2i - 4r. \quad (31.14)$$

Мы видим выходящий из капли луч под углом

$$\theta = \pi - \theta_0 = -2i + 4r. \quad (31.15)$$

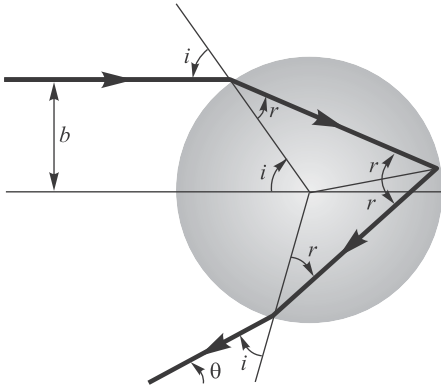
Углы  $i$  и  $r$  связаны законом Снеллиуса:

$$\frac{\sin i}{\sin r} = n, \quad (31.16)$$

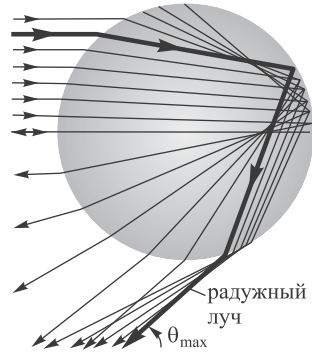
где  $n$  — показатель преломления вещества капли (воды). Угол падения зависит от **прицельного расстояния**  $b$  — расстояния между направлением падающего луча и параллельной ему линией, проходящей через центр капли:

$$\sin i = \frac{b}{R}, \quad (31.17)$$

<sup>4</sup>Л. И. Пономарев. Числа радуги. Химия и жизнь, 1981, № 10, С. 44–50.



**Рис. 31.10.** Ход в дождевой капле одного из солнечных лучей, приводящих к образованию радуги



**Рис. 31.11.** Большая часть солнечных лучей, рассеянных дождевой каплей, выходит из нее в преимущественном направлении  $\theta_{\max}$

где  $R$  — радиус капли. Вводя безразмерный параметр  $\rho = b/R$  и группируя соотношения (31.15) — (31.17), приходим к уравнению

$$\theta = 4 \arcsin(\rho/n) - 2 \arcsin \rho. \quad (31.18)$$

Декарт провел численные расчеты и обнаружил, что хотя падающие на каплю лучи распределены равномерно по прицельному параметру  $\rho$ , выходящие из капли лучи распределены по углам неравномерно: большая их часть сосредоточена вблизи угла  $\theta_{\max}$  (рис. 31.11). Сейчас такой расчет может провести любой студент: достаточно найти экстремум функции  $\theta(\rho)$ , определенной уравнением (31.18):

$$\frac{d\theta}{d\rho} = \frac{2}{\sqrt{1-\rho^2}} - \frac{4}{\sqrt{n^2-\rho^2}} = 0. \quad (31.19)$$

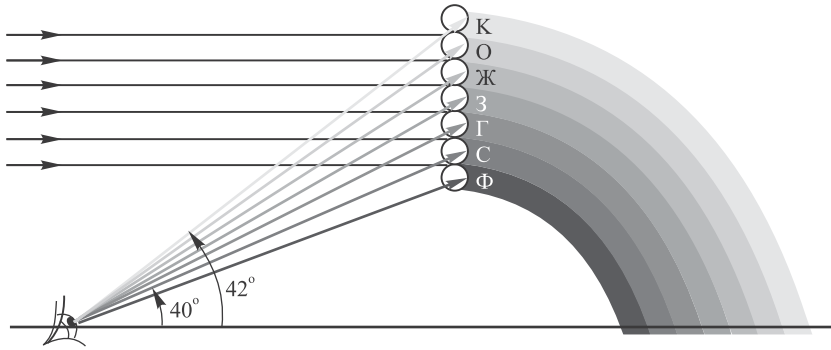
Отсюда находим прицельное расстояние  $\rho_{\max}$ , соответствующее углу, при котором сгущаются выходящие из капли лучи:

$$\rho_{\max} = \sqrt{\frac{4-n^2}{3}}. \quad (31.20)$$

Угол  $\theta_{\max}$ , под которым к нам приходят лучи от радуги, получается подстановкой  $\rho_{\max}$  в (31.18).

Ньютон повторил опыты де Доминиса и Декарта со стеклянными шарами, наполненными водой (рис. 31.12). Вот его описание:

«Это объяснение радуги подтверждается далее известным опытом (сделанным Антонием де Доминисом и Декартом), заключающимся в подвешивании где-нибудь на солнечном свете стеклянного шара, наполненного водой; шар рассматривается в таком положении, чтобы лучи, идущие от него к глазу, составляли с солнечными лучами угол в 42 или 50 градусов. Ибо если угол будет около 42 или 43 градусов, то зритель увидит полный красный цвет на стороне шара, противоположной солнцу; если этот угол будет меньше (например, при опускании шара), то будут видны другие



**Рис. 31.12.** Схема опыта Ньютона со стеклянным шаром, наполненным водой: по мере подъема шара его нижняя половина окрашивается во все цвета радуги

цвета: последовательно желтый, зеленый и синий с той же стороны шара. Но если угол будет около 50 градусов (например, при подъеме шара), то со стороны шара, обращенной к солнцу, покажется красный цвет; если же угол будет делаться больше (например, при подъеме шара), то красный будет переходить последовательно в другие цвета: желтый, зеленый и синий. Я получал то же, оставляя шар в покое, но поднимая или опуская глаз, или двигая его иным способом, так, чтобы получить угол правильной величины»<sup>5</sup>.

Показатель преломления воды для красных лучей равен  $n = 1,331$ , откуда следует  $\theta_{\max} = 42^\circ 22'$ . Для фиолетовых лучей  $n = 1,343$ , чему соответствует значение угла  $\theta_{\max} = 40^\circ 39'$ . Следовательно, угловая ширина радуги примерно равна двум градусам. Таким образом, дождевая капля рассеивает солнечные лучи в конус с центром в точке местонахождения наблюдателя и углом раствора  $84^\circ$ . Вблизи поверхности конуса интенсивность света резко возрастает, за его пределами находится область геометрической тени, куда не попадают рассеянные каплей лучи.

Наблюдения показали также, что свет от радуги сильно поляризован (в соотношении примерно 21:1). Это также может быть объяснено в рамках теории Декарта. Для оптимального угла относительное прицельное расстояние (синус угла падения) дается формулой (31.20). На основе закона Снеллиуса найдем отсюда синус, а затем и тангенс угла преломления:

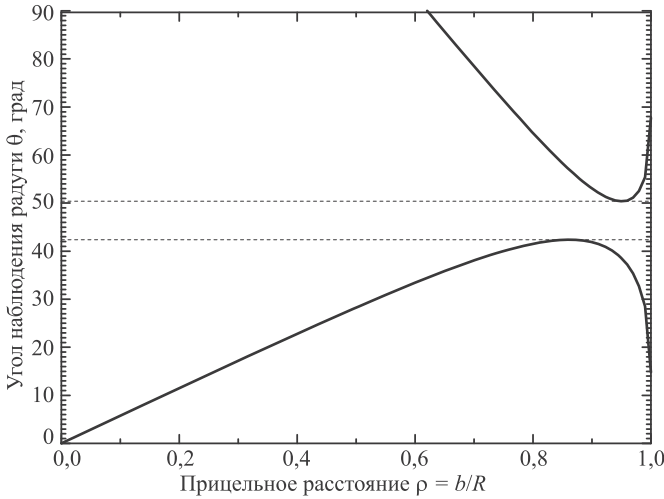
$$\sin r = \frac{1}{n} \sqrt{\frac{4(n^2 - 1)}{3}}, \quad \operatorname{tg} r = \sqrt{\frac{4 - n^2}{4(n^2 - 1)}}. \quad (31.21)$$

Отсюда следует, например, что при  $n = 1,343$  (красный цвет) угол преломления  $r \approx 50^\circ$  близок к углу Брюстера  $i_B = \operatorname{arctg} n \approx 53^\circ$ , при котором свет радуги был бы полностью поляризован.

При расчете радуги второго порядка следует учесть, что при отражениях луч поворачивается на угол  $\pi - 2r$  дважды, так что вместо уравнения (31.18) к приходим к уравнению

$$\theta_0 = 2\pi + 2 \operatorname{arcsin} \rho - 6 \operatorname{arcsin}(\rho/n). \quad (31.22)$$

<sup>5</sup> И. Ньютон. Оптика. М.: Гостехиздат, 1954. С. 132.



**Рис. 31.13.** Зависимость от прицельного расстояния угла наблюдения радуги первого (нижняя кривая) и второго (верхняя кривая) порядков для красного цвета ( $n = 1,331$ ); точки экстремумов соответствуют углам  $\theta_{\max} = 42,37^\circ$  и  $\theta_{\max} \approx 50,37^\circ$  (показаны пунктирными линиями), промежуток между ними — темная александрова полоса

Связь угла поворота  $\theta_0$  с углом наблюдения радуги  $\theta$  в этом случае дается соотношением  $\theta_0 = \pi + \theta$ , так что

$$\theta = \pi + 2 \arcsin \rho - 6 \arcsin(\rho/n). \quad (31.23)$$

Отсюда вытекает, что из второй радуги красный луч выходит под углом  $\theta_{\max} = 50^\circ 22'$ , а фиолетовый — под углом  $\theta_{\max} = 53^\circ 29'$ . Стало быть, ширина радуги второго порядка составляет три градуса, а ширина александровой полосы (угловое расстояние между красными полосами в радугах первого и второго порядков) — восемь градусов.

Графики зависимости угла наблюдения от прицельного расстояния для радуг первого и второго порядка показаны на рис. 31.13. Как видно, в область геометрической тени  $42^\circ < \theta < 50^\circ$  световые лучи вообще не попадают, что объясняет наличие более темной александровой полосы между двумя радугами. В области же  $\theta < 42^\circ$  и  $\theta > 50^\circ$  какая-то часть лучей попадает, так что яркость областей внутри этих конусов соответствует яркости основного фона неба.

Следует отметить, что простая геометрическая теория радуги, основанная на явлении преломления, дает лишь элементарное описание этого феномена. Например, во внутренней части первой радуги и во внешней части второй вблизи их вершин (т. е. в светлых областях вблизи фиолетовых дуг) наблюдаются так называемые *дополнительные дуги*, окрашенные в фиолетово-розовые и нежно-зеленые цвета. Их происхождение связано с явлением, о котором не знали ни Декарт, ни Ньютон. Как видно из рисунка 31.13, во внутренние области радуг попадают лучи, падающие на каплю с двумя разными прицельными расстояниями. Возникает интерференция, совершенно аналогичная интерференции в двухщелевом опыте Юнга. Неудивительно поэтому, что именно Т. Юнг предложил (1803) интерференционную картину (по крайней мере, качественную) распределения интенсивности радуги с учетом этого явления. В отличие от экстремального угла  $\theta_{\max}$  положение интерференционных пиков и минимумов зависит от размера капли. В больших каплях разность хода лучей сильно зависит от

прицельного параметра, дополнительные дуги находятся на малом расстоянии друг от друга, перекрываются и плохо наблюдаются при диаметрах капель больше 1–2 мм. При малых каплях (с диаметром, меньшим 0,02 мм) радуга бледнеет и превращается в так называемую *белую радугу*.

В интерференционной теории Юнга есть такой же резкий переход между внутренними областями радуги и александровой полосой. Но в волновой теории света таких резких границ между светом и тенью быть не должно. Его нет и в дифракционной теории радуги, построенной английским математиком и астрономом Дж. Эйри (1838). Он использовал принцип Гюйгенса, рассматривая фронт волны в капле как источник вторичных волн. В его теории из-за дифракции часть света будет попадать также и в область геометрической тени (т.е. в александрову полосу). Дифракционная теория радуги позволила количественно рассчитать распределение интенсивностей в радуге и изменение цвета дополнительных дуг в зависимости от размеров капель. Теперь по виду радуги метеорологи могут определить размеры дождевых капель. Заметим, что дифракционная теория предсказывает несколько меньшее значение угла рассеяния, нежели теории Декарта и Юнга.

Дефект теории Эйри состоял в том, что приходилось угадывать амплитуду волны вдоль произвольно выбранного начального фронта. Полностью корректное решение задачи о рассеянии электромагнитной волны на сферическом диэлектрическом шарике дали в рамках теории Максвелла немецкий физик Г. Ми и будущий лауреат Нобелевской премии по химии голландец П. Дебай (1908). Решения оказались очень сложными и малоприменимыми для сравнения с наблюдаемыми фактами. Пришлось разработать мощные математические методы, и в настоящее время в теории радуги не осталось как будто белых пятен. Точные решения позволили определить области параметров, где более простая теория Эйри приводит к совершенно неправильным предсказаниям.

Но история на этом не закончилась. Радужное рассеяние наблюдается также и в механике — например, при рассеянии легких ионов на атомных ядрах или электронов на цепочке атомов в кристалле (где уже надо применять квантовую механику).

## Элементарная теория дисперсии

В 1860 г. французский физик Ф. П. Лё Ру обнаружил явление, которое он назвал *аномальной дисперсией*: в парах иода синие лучи преломлялись меньше, чем красные. Детальные исследования аномальной дисперсии провел немецкий ученый А. Кундт, учитель В. Рентгена и П. Н. Лебедева. Он обнаружил, что аномальное поведение дисперсии  $dn/d\lambda > 0$  ( $dn/d\omega < 0$ ) наблюдается у веществ в том диапазоне частот, где они сильно поглощают свет.

Зависимость показателя преломления вещества от длины электромагнитной волны объясняется на основе теории вынужденных колебаний. Строго говоря, движение электронов в атоме подчиняется законам квантовой механики. Однако для качественного понимания оптических явлений можно ограничиться представлением об электронах, связанных в атоме упругой силой. При отклонении от равновесного положения такие электроны начинают колебаться, постепенно теряя энергию на излучение электромагнитных волн или передавая свою энергию узлам решетки и нагревая вещество. В результате этого колебания будут затухающими.

При прохождении через вещество электромагнитная волна воздействует на каждый электрон с силой Лоренца:

$$\vec{F} = -e\vec{E} - e[\vec{v}, \vec{B}] = -e\vec{E} - e\mu_0[\vec{v}, \vec{H}], \quad (31.24)$$

где  $\vec{v}$  — скорость колеблющегося электрона. В электромагнитной волне отношение напряженностей магнитного и электрического полей равно:

$$\frac{H}{E} = \sqrt{\frac{\varepsilon_0}{\mu_0}}. \quad (31.25)$$

Поэтому нетрудно оценить отношение электрической и магнитной сил, действующих на электрон:

$$\frac{F_{\text{эл}}}{F_{\text{м}}} = \frac{eE}{e\mu_0vH} = \frac{1}{v\sqrt{\varepsilon_0\mu_0}} = \frac{c}{v}. \quad (31.26)$$

Электроны в веществе движутся со скоростями, много меньшими скорости света в вакууме:  $v/c < 10^{-2}$ . Таким образом, можно считать, что при прохождении через вещество электромагнитной волны на каждый электрон действует только электрическая сила

$$F = -eE(t), \quad E(t) = E_0e^{i\omega t}, \quad (31.27)$$

где  $E_0$  — амплитуда напряженности электрического поля в световой волне. Запишем уравнение движения электрона:

$$m\ddot{x} = -m\omega_0^2x - m\gamma\dot{x} - eE_0e^{i\omega t}, \quad (31.28)$$

где  $-m\omega_0^2x$  — аналог упругой силы, возвращающей электрон в положение равновесия ( $\omega_0$  — собственная частота колебаний электрона в атоме), а  $-m\gamma\dot{x}$  — аналог силы трения, направленной против скорости электрона ( $\gamma$  — постоянная затухания).

Решение такого дифференциального неоднородного уравнения мы уже получили в разделе 24.7:

$$x(t) = -\frac{e}{m} \cdot \frac{1}{\omega_0^2 - \omega^2 + i\gamma\omega} E(t). \quad (31.29)$$

Следовательно, смещение электрона из положения равновесия пропорционально напряженности электрического поля. Смещениями ядер из положения равновесия можно пренебречь, так как массы ядер велики по сравнению с массой электрона.

Атом со смещенным электроном приобретает дипольный момент  $p(t) = -ex(t)$ . Если в единице объема находится  $N$  атомов, которые поляризуются, то дипольный момент единицы объема (поляризация среды) равен  $P(t) = Np(t)$ , т. е.

$$P(t) = \frac{e^2}{m} \cdot \frac{N}{\omega_0^2 - \omega^2 + i\gamma\omega} E(t). \quad (31.30)$$

В реальных средах возможны разные типы колебаний различных групп электронов (мы будем обозначать их индексом  $k$ ). У этих колебаний могут быть разные заряды  $e_k$ , массы  $m_k$  и собственные частоты  $\omega_{0k}$ , а число  $N_k$  атомов, участвующих в данном типе колебаний, пропорционально полному числу  $N_0$  атомов в единице объема:  $N_k = f_k N_0$ . Безразмерный коэффициент  $f_k$  обычно называют *силой осциллятора*, подчеркивая, что его величина указывает на долю участия данных колебаний в явлении дисперсии. Все эти колебания дают вклад в поляризуемость среды, так что (31.30) заменится на сумму

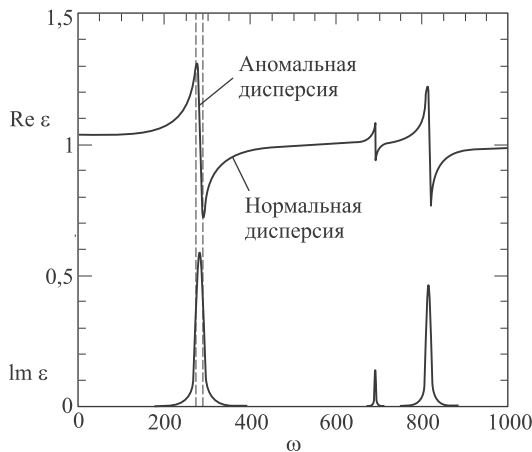
$$P(t) = E(t)N_0 \sum_k \frac{e_k^2}{m_k} \cdot \frac{f_k}{\omega_{0k}^2 - \omega^2 + i\gamma_k \omega}. \quad (31.31)$$

В то же время  $\vec{P}(t) = \chi_{\varepsilon 0} \vec{E}(t)$ , где  $\chi$  — диэлектрическая восприимчивость вещества, которая связана с диэлектрической проницаемостью  $\varepsilon$  соотношением  $\varepsilon = n^2 = 1 + \chi$  (см. т. 1, разд. 17.3). В результате получаем выражение для диэлектрической проницаемости (квадрата показателя преломления) вещества

$$\varepsilon = n^2 = 1 + \frac{N_0}{\varepsilon_0} \sum_k \frac{e_k^2}{m_k} \cdot \frac{f_k}{\omega_{0k}^2 - \omega^2 + i\gamma_k \omega}. \quad (31.32)$$

Коэффициент затухания обычно мал по сравнению с собственными частотами колебаний в атоме: типичные значения  $\gamma_k \approx 10^9$  Гц,  $\omega_{0k} \approx 10^{15}$  Гц.

Функция  $\varepsilon = n^2$  имеет вид, изображенный на рис. 31.14. Если частота падающей волны меньше наименьшей из резонансных частот, все члены суммы в формуле (31.32) положительны и показатель преломления больше единицы. При прохождении через первую резонансную частоту соответствующее слагаемое резко возрастает, а затем падает до нуля и становится отрицательным



**Рис. 31.14.** Зависимость диэлектрической проницаемости  $\varepsilon$  от частоты электромагнитной волны (график с сайта физического факультета университета Дьюка, Сев. Каролина, США)



( $n < 1$ ). При дальнейшем повышении частоты это слагаемое медленно растет, приближаясь к нулю. Такое поведение кривой наблюдается при прохождении каждой из резонансных частот; на рис. 31.14 показаны три таких резонанса. При высоких частотах, превышающих все резонансные частоты, показатель преломления стремится к единице снизу.

Показатель преломления ассоциируется с действительной частью  $\text{Re} \varepsilon$  диэлектрической проницаемости. С ее мнимой частью, показанной на нижней половине рисунка, связано резонансное поглощение света. В области пика  $\text{Im} \varepsilon$  быстро растет и достигает максимума в точке, где  $\text{Re} \varepsilon = 0$ . Затем мнимая часть диэлектрической проницаемости быстро падает до нуля. Таким образом, в области резонанса доминирует мнимая часть диэлектрической проницаемости. Появление мнимой части у показателя преломления означает появление мнимой части у волнового вектора, а это означает поглощение волны, как мы увидим ниже в разделах 32.8 (о полном внутреннем отражении) и 32.9 (о затухании волн в металле и скин-эффекте).

### 31.3 Поглощение света

Световая волна несет поток энергии электромагнитного поля. При прохождении света через вещество происходит потеря энергии из-за превращения ее в различные формы внутренней энергии вещества или энергию вторичного излучения, которое может отличаться от первичного спектральным составом и направлением распространения. Поглощение света может приводить к нагреванию вещества или возбуждению атомов и молекул, к фотохимическим процессам и т. п. Свободные электроны в металлах поглощают всю энергию световой волны, так что металлы практически непрозрачны для света (волн видимого диапазона). Тонкие слои металла могут быть, однако, прозрачными для электромагнитного излучения других длин волн.

Если световая волна с интенсивностью в данной точке  $I(x)$  проходит через слой толщиной  $dx$ , то ее интенсивность уменьшается на величину, пропорциональную толщине слоя:

$$dI = -I(x)\kappa dx, \quad (31.33)$$

где  $\kappa$  — *коэффициент поглощения*, зависящий от свойств поглощающего вещества. Решаем полученное дифференциальное уравнение:

$$\frac{dI}{I} = -\kappa dx, \quad \text{откуда} \quad \int_{I_0}^I \frac{dI}{I} = -\kappa \int_0^l dx. \quad (31.34)$$

В результате получаем

$$I = I_0 e^{-\kappa l}, \quad (31.35)$$

где  $I_0$  — падающий световой поток,  $l$  — толщина поглощающего слоя вещества. Это соотношение называется *законом Бугера*. Он справедлив для монохроматического света. Соотношение (31.35) можно записать в виде

$$I = I_0 e^{-l/l_0}, \quad (31.36)$$

где  $l_0 = 1/\kappa$ . Отсюда вытекает физический смысл  $\kappa$ : коэффициент поглощения обратен толщине слоя вещества, при прохождении которого интенсивность света уменьшается в  $e \approx 2,72$  раза. Этот закон носит приближенный характер. В частности, он не учитывает зависимость показателя поглощения от длины волны и не работает в области нелинейной оптики.

## 31.4 Рассеяние света

Мы наблюдаем красное солнце на закате и голубое небо в ясный день. Эти явления можно объяснить рассеянием света на молекулах атмосферы. Переменное электрическое поле световой волны возбуждает в облучаемой молекуле воздуха вынужденные колебания, вследствие чего молекула приобретает изменяющийся во времени электрический дипольный момент и становится источником собственного излучения, частота которого равна частоте облучающего диполь света. Колеблющиеся электроны возбуждают вторичные волны, распространяющиеся по всем направлениям, а частота рассеянного света равна частоте света в первичном пучке.

Согласно теории излучения на основе уравнений Максвелла мощность вторичного излучения пропорциональна квадрату ускорения заряженной частицы (см. разд. 32.10). Если электроны под действием световой волны колеблются по закону  $x(t) = x_0 \cos(\omega t + \alpha)$ , то их ускорение  $\ddot{x}(t)$  пропорционально квадрату частоты  $\omega$ . Соответственно мощность вторичного излучения пропорциональна четвертой степени частоты. Поэтому и интенсивность рассеянного света пропорциональна частоте в четвертой степени или обратно пропорциональна четвертой степени длины световой волны:

$$I_{\text{рас}} \sim \omega^4 \sim \frac{1}{\lambda^4}. \quad (31.37)$$

Это соотношение называется **законом Рэлея**. Следовательно, красный и оранжевый свет рассеиваются гораздо слабее, чем голубой и фиолетовый, и поэтому небо в ясный день выглядит голубым. На закате солнечные лучи проходят максимальный слой атмосферы. Значительная часть голубого излучения рассеивается и не попадает в глаз наблюдателя. Поэтому прошедший луч, достигающий поверхности Земли, лишается голубой и фиолетовой составляющих и кажется красноватым.

Следует отметить, что закон Рэлея имеет место только в случае, когда рассеивающие объекты оказываются меньше длины световой волны. Это условие выполняется для молекул кислорода и азота, имеющих размеры около 0,2 нм. Но облака содержат капельки воды и кристаллы льда, размеры которых значительно превышают длину волны. Такие частицы рассеивают свет почти равномерно на всех частотах, поэтому облака кажутся белыми.

## Контрольные вопросы

1. Что такое линейно-поляризованная электромагнитная волна и волна с круговой поляризацией?

2. Чем различаются естественный свет и поляризованный?
3. Сформулируйте закон Малюса.
4. В чем состоит закон Брюстера? Как его можно объяснить с позиции теории излучения электромагнитных волн?
5. Что такое дисперсия света?
6. Нарисуйте кривую зависимости показателя преломления вещества от частоты электромагнитной волны и объясните качественно ее ход.
7. По какому закону происходит изменение интенсивности светового пучка при прохождении через слой вещества?
8. Сформулируйте закон Рэлея для рассеяния света. Объясните на его основе, почему небо кажется голубым.
9. Какого цвета было бы небо при отсутствии атмосферы?
10. Почему солнце в зените кажется желтым, а на закате — красным? Почему красный цвет выбран на транспорте как сигнал опасности?

## Глава 32

# Электромагнитные волны

Волновая теория света берет свое начало от Х. Гюйгенса. Она основана на предположении, что свет представляет собой волновой процесс. Какая среда при этом колеблется, во времена Гюйгенса было непонятно. Эту среду — мировой эфир — представляли себе механистически, а его колебания — световые волны — подобными колебаниям твердого тела. Исходя из таких самых общих представлений о волновых процессах удалось описать явления интерференции, дифракции, поляризации, дисперсии. Волновая теория света была полностью подтверждена в многочисленных экспериментах, хотя применение представлений о колебаниях упругого эфира иногда наталкивалось на трудности.

После вывода Дж. К. Максвеллом (1873) уравнений классической электродинамики стало понятно, что переменные электрическое и магнитное поля могут существовать и распространяться в пустом пространстве, в отсутствие зарядов и токов. Максвелл первым заметил, что скорость распространения электромагнитных волн близка к скорости света, и это навело физиков на мысль, что свет — одна из разновидностей электромагнитных волн. Их существование было подтверждено экспериментально Г. Герцем (1887–1888). В конечном итоге последовательное и точное описание законов волновой оптики было дано в рамках электродинамики Максвелла.

### 32.1 Усреднение микроскопических полей

Уравнения Максвелла в вакууме относятся к самым фундаментальным и точным законам физики. Однако для описания электромагнитных полей в какой-то среде необходимо учитывать свойства этой среды. Поскольку вещество состоит из атомов и молекул, для их описания нужна квантовая механика. Но даже если ограничиваться классическими моделями, то вещество имеет очень разные свойства — есть проводники, диэлектрики, полупроводники, плазма, ферромагнетики, сегнетоэлектрики и т. д. и т. п. В веществе существуют разные механизмы проводимости. Задача выведения уравнений электродинамики в средах кажется неразрешимой. И все же можно сделать достаточно общие предположения и описывать вещество как сплошную среду с некоторыми средними характеристиками и какими-то феноменологическими параметрами, отражающими его свойства.

*Сплошная среда* — модель более или менее однородной системы с очень большим числом частиц. Такая модель используется в механике — теория упругости, гидро- и аэродинамика, теория плазмы. Такая же модель применяется для описания электромагнитных полей в веществе.

Все электромагнитные явления можно описать уравнениями Максвелла (см. т. 1, гл. 23), в которых мы на этом этапе слегка изменим обозначения для напряженности электрического поля и магнитной индукции:

$$\begin{aligned} \operatorname{div} \vec{e} &= \frac{\rho'_e}{\varepsilon_0}, & \operatorname{rot} \vec{e} &= -\frac{\partial \vec{b}}{\partial t}, \\ \operatorname{div} \vec{b} &= 0, & \operatorname{rot} \vec{b} &= \mu_0 \left( \vec{j}' + \varepsilon_0 \frac{\partial \vec{e}}{\partial t} \right). \end{aligned} \quad (32.1)$$

Эти уравнения — обычные уравнения Максвелла для полей в вакууме — описывают микроскопические электрическое  $\vec{e}$  и магнитное  $\vec{b}$  поля, создаваемые полными «микроскопическими» плотностями заряда и тока (помечены штрихами).

В принципе мы умеем решать эти уравнения для полей, создаваемых в вакууме определенными распределениями зарядов и токов. Но при переходе к рассмотрению электромагнитных полей в веществе мы имеем дело с огромным числом зарядов (порядка  $10^{23}$ ), находящихся к тому же в постоянном тепловом движении и квантовом «дрожании», с которым мы познакомимся в следующей главе курса. Соответственно поля в веществе сильно меняются на расстояниях порядка размеров атомов или молекул (для оценки возьмем 1 нм). Таким образом, в кубическом миллиметре вещества содержится  $10^{18}$  атомов или молекул. Это означает, что на таких малых расстояниях микроскопические флуктуации полностью усредняются. Но нам и не нужна подробная информация о микроскопических полях — достаточно сведений о полях, усредненных по небольшим макроскопическим объемам.

Для средних значений полей в веществе мы используем прежние обозначения:

$$\vec{E} = \langle \vec{e} \rangle, \quad \vec{B} = \langle \vec{b} \rangle. \quad (32.2)$$

Усредняя (32.1) по небольшим, но макроскопическим объемам, получаем уравнения для полей в среде

$$\begin{aligned} \operatorname{div} \vec{E} &= \frac{\langle \rho'_e \rangle}{\varepsilon_0}, & \operatorname{rot} \vec{E} &= -\frac{\partial \vec{B}}{\partial t}, \\ \operatorname{div} \vec{B} &= 0, & \operatorname{rot} \vec{B} &= \mu_0 \left( \langle \vec{j}' \rangle + \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \right). \end{aligned} \quad (32.3)$$

Микроскопическая плотность заряда состоит из двух компонентов — плотности  $\rho_{\text{ст}}$  сторонних зарядов, внесенных в вещество извне, и плотности  $\rho_{\text{вн}}$  связанных (внутренних) зарядов, входящих в структурные элементы вещества — атомы и молекулы:  $\rho'_e = \rho_{\text{ст}} + \rho_{\text{вн}}$ . Среднее значение плотности сторонних зарядов и есть та плотность, с которой мы привыкли работать:  $\langle \rho_{\text{ст}} \rangle = \rho_e$ . Среднюю плотность внутренних зарядов мы отметим значком тильды:  $\langle \rho_{\text{вн}} \rangle = \tilde{\rho}_e$ . Среднее значение микроскопической плотности заряда мы можем тогда записать в виде

$$\langle \rho'_e \rangle = \rho_e + \tilde{\rho}_e. \quad (32.4)$$

Так как вещество (диэлектрик) в целом электрически нейтрально, то интеграл от средней плотности внутренних зарядов по всему объему тела равен нулю:

$$\int_V \tilde{\rho}_e dV = 0. \quad (32.5)$$

Значит, среднюю плотность внутренних зарядов можно представить как дивергенцию некоторого вектора  $\vec{P}$ , причем последний равен нулю вне тела:

$$\tilde{\rho}_e = -\operatorname{div} \vec{P}. \quad (32.6)$$

Можно вывести связь полного дипольного момента  $\vec{p}$  внутренних зарядов с пока формально введенным вектором  $\vec{P}$ :

$$\vec{p} = \int_V \vec{r} \tilde{\rho}_e dV = \int_V \vec{P} dV. \quad (32.7)$$

Таким образом,  $\vec{P}$  — это дипольный момент внутренних зарядов в единице объема диэлектрика, т. е. вектор поляризации, с которым мы имели дело в главе 17.

Первое уравнение Максвелла, из которого следует закон Кулона, модифицируется в среде следующим образом:

$$\operatorname{div} \vec{E} = \frac{1}{\varepsilon_0} (\rho_e - \operatorname{div} \vec{P}). \quad (32.8)$$

Вводя также уже знакомый нам вектор электрического смещения  $\vec{D} = \varepsilon_0 \vec{E} + \vec{P}$ , записываем это уравнение Максвелла в виде

$$\operatorname{div} \vec{D} = \rho_e. \quad (32.9)$$

Напомним, что  $\rho_e$  — это плотность сторонних зарядов, внесенных в наш образец.

Чтобы уравнения Максвелла по-прежнему составляли полную систему уравнений электродинамики в веществе, необходимо установить связь между напряженностью электрического поля  $\vec{E}$  и вектором электрического смещения  $\vec{D}$ . В слабых полях достаточно ограничиться первым членом разложения  $\vec{D}$  по степеням  $\vec{E}$ . Тогда  $\vec{D}$  зависит от  $\vec{E}$  по линейному закону, который для изотропных и однородных сред имеет вид  $\vec{D} = \varepsilon_0 \varepsilon \vec{E}$ , и уравнение (32.9) принимает вид

$$\operatorname{div} \vec{E} = \frac{\rho_e}{\varepsilon_0 \varepsilon}. \quad (32.10)$$

Теперь надо усреднить последнее уравнение Максвелла, куда входит среднее значение плотности тока в веществе. Пусть  $\vec{j}$  — средняя плотность тока, создаваемого сторонними зарядами (среднюю плотность которых мы обозначили выше как  $\rho_e$ ). Уравнение непрерывности (закон сохранения заряда) имеет вид

$$\frac{\partial \rho_e}{\partial t} + \operatorname{div} \vec{j} = 0. \quad (32.11)$$

Производную плотности зарядов можно выразить через вектор электрического смещения с помощью уравнения Максвелла (32.9). Имеем тогда

$$\operatorname{div} \left( \frac{\partial \vec{D}}{\partial t} + \vec{j} \right) = 0. \quad (32.12)$$

Равенство нулю дивергенции вектора означает, что его можно представить как ротор какого-то другого вектора:

$$\operatorname{rot} \vec{H} = \vec{j} + \frac{\partial \vec{D}}{\partial t}. \quad (32.13)$$

Для выяснения физического смысла этого вектора следует разложить среднее значение микроскопического тока по степеням полей и оставить лишь линейные члены. Мы приведем лишь окончательный результат:

$$\langle \vec{j}' \rangle = \sigma \vec{E} + \frac{\partial \vec{P}}{\partial t} + \vec{J}, \quad (32.14)$$

где  $\vec{J}$  — средний магнитный момент единицы объема тела, вызываемый движущимися связанными зарядами, т. е. вектор намагничивания, с которым мы уже имели дело (см. т. 1, гл. 21). Подставляя (32.14) в последнее уравнение Максвелла в системе (32.3), получаем

$$\operatorname{rot} \vec{B} = \mu_0 \left( \sigma \vec{E} + \frac{\partial \vec{D}}{\partial t} + \vec{J} \right). \quad (32.15)$$

Сравнивая (32.13) и (32.15), находим уравнения связи:

$$\vec{j} = \sigma \vec{E} \quad (32.16)$$

и

$$\vec{H} = \frac{1}{\mu_0} \vec{B} - \vec{J}. \quad (32.17)$$

Таким образом, вектор  $\vec{H}$  — введенная ранее напряженность магнитного поля, а  $\sigma$  — удельная проводимость среды.

## 32.2 Уравнения Максвелла для сплошных сред

Выпишем снова полученные нами уравнения Максвелла в среде

$$\boxed{\begin{aligned} \operatorname{div} \vec{D} &= \rho_e, & \operatorname{rot} \vec{E} &= -\frac{\partial \vec{B}}{\partial t}, \\ \operatorname{div} \vec{B} &= 0, & \operatorname{rot} \vec{H} &= \vec{j} + \frac{\partial \vec{D}}{\partial t} \end{aligned}} \quad (32.18)$$

и уравнения связи

$$\boxed{\begin{aligned} \vec{D} &= \varepsilon_0 \vec{E} + \vec{P}, \\ \vec{H} &= \frac{1}{\mu_0} \vec{B} - \vec{J}, \\ \vec{j} &= \sigma \vec{E}. \end{aligned}} \quad (32.19)$$

В случае однородных и изотропных сред уравнения связи имеют вид

$$\boxed{\begin{aligned} \vec{D} &= \varepsilon_0 \varepsilon \vec{E}, & \vec{B} &= \mu_0 \mu \vec{H}; \\ \vec{P} &= \varepsilon_0 (\varepsilon - 1) \vec{E}, & \vec{J} &= \frac{\mu - 1}{\mu} \frac{\vec{B}}{\mu_0}. \end{aligned}} \quad (32.20)$$

Плотность сторонних зарядов и ток проводимости связаны уравнением непрерывности

$$\frac{\partial \rho_e}{\partial t} + \operatorname{div} \vec{j} = 0. \quad (32.21)$$

Теоремы о дивергенции и роторе позволяют записать систему уравнений (32.18) в интегральной форме:

$$\boxed{\begin{aligned} \oint_S \vec{D} \cdot d\vec{S} &= \int_V \rho_e dV, & \oint_L \vec{E} \cdot d\vec{l} &= - \int_S \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S}, \\ \oint_S \vec{B} \cdot d\vec{S} &= 0, & \oint_L \vec{H} \cdot d\vec{l} &= \int_S \left( \vec{j} + \frac{\partial \vec{D}}{\partial t} \right) \cdot d\vec{S}. \end{aligned}} \quad (32.22)$$

В отличие от уравнений Максвелла в вакууме, уравнения электромагнитного поля в среде не имеют такого фундаментального характера из-за ограниченности области применения уравнений связи, что следует учитывать при рассмотрении конкретных систем. При выводе уравнений электродинамики сплошных сред были сделаны довольно сильные предположения.

- Среда предполагалась изотропной, поэтому ее можно было охарактеризовать скалярными коэффициентами в разложении плотности тока по степеням полей. Если это не так, то возникнут тензоры диэлектрической и магнитной проницаемости.
- Линейная связь поляризации с напряженностью электрического поля и намагниченности с магнитной индукцией предполагает слабость полей, но не только: в ряде веществ (ферромагнетики, антиферромагнетики, сегнетоэлектрики) зависимость  $\vec{P}(\vec{E})$  и  $\vec{J}(\vec{B})$  нелинейна и неоднозначна (примером может служить петля гистерезиса). Другой пример — сверхпроводники, внутрь которых не проникает магнитное поле.
- Предполагается, что частоты  $\omega$  полей не слишком высоки — много меньше частот  $\omega_0$  внутриатомных колебаний:  $\omega \ll \omega_0 \approx 10^{18}$  Гц. При высоких частотах  $\omega \approx \omega_0$  (рентгеновский диапазон) колебания сторонних и связанных зарядов уже неотличимы. Одновременно это означает, что длина электромагнитной волны в этом случае порядка размера атома, и усреднение по макроскопическому объему, содержащему много атомов, теряет смысл.
- Даже при  $\omega \ll \omega_0$  уравнения связи  $P \propto E$  и  $J \propto B$  могут оказаться неприменимыми. Частота поля должна быть много меньше обратного времени релаксации системы, иначе поляризация и намагниченность будут определяться не мгновенным значением электрического и магнитного полей, а всей предысторией образца.
- Если длина электромагнитной волны порядка размера неоднородностей вещества, то возникает пространственная дисперсия: вклад в поляризацию среды дают заряды, находившиеся в соседних точках в предшествующие моменты времени.



### 32.3 Волновое уравнение

Запишем уравнения Максвелла в дифференциальной форме для случая, когда отсутствуют электрические заряды ( $\rho = 0$ ) и токи ( $\vec{j} = 0$ ):

$$\begin{aligned} \operatorname{div} \vec{E} &= 0, & \operatorname{rot} \vec{E} &= -\frac{\partial \vec{B}}{\partial t}, \\ \operatorname{div} \vec{H} &= 0, & \operatorname{rot} \vec{H} &= \frac{\partial \vec{D}}{\partial t}, \end{aligned} \quad (32.23)$$

где  $\vec{D} = \varepsilon \varepsilon_0 \vec{E}$  и  $\vec{B} = \mu \mu_0 \vec{H}$ . Величины  $\varepsilon_0$  и  $\mu_0$  — электрическая и магнитная постоянные соответственно, которые связаны со скоростью света в вакууме ( $c$ ) соотношением  $\varepsilon_0 \mu_0 = 1/c^2$ . Постоянные  $\varepsilon$  и  $\mu$  характеризуют электрические и магнитные свойства среды.

При отсутствии зарядов и токов невозможно существование статических электрического и магнитного полей. Однако переменное электрическое поле возбуждает магнитное поле и, наоборот, переменное магнитное поле создает переменное электрическое поле. Поэтому имеются решения уравнений Максвелла в вакууме, при отсутствии зарядов и токов, где электрические и магнитные поля оказываются неразрывно связанными друг с другом. В теории Максвелла впервые были объединены два фундаментальных взаимодействия — электрическое и магнитное, которые ранее считались независимыми. Поэтому мы говорим теперь об *электромагнитном поле*.

Колебательный процесс в электрическом контуре или вибраторе (проводнике, где возбужден переменный электрический ток) сопровождается изменением окружающего его поля. Изменения, происходящие в окружающем пространстве, распространяются от точки к точке с определенной скоростью, т.е. колебательный контур излучает в окружающее его пространство энергию электромагнитного поля. Распространяющееся в пространстве электромагнитное поле, в котором электрическая напряженность и магнитная индукция изменяются по периодическому закону, называется *электромагнитной волной*. Получим из уравнений Максвелла волновые уравнения для векторов  $\vec{E}$  и  $\vec{H}$ .

Напомним, что ротор ( $\operatorname{rot}$ ) и дивергенция ( $\operatorname{div}$ ) — операции дифференцирования, производимые по определенным правилам над векторными полями. Возьмем ротор от обеих частей уравнения  $\operatorname{rot} \vec{E} = -\partial \vec{B} / \partial t$ . При этом воспользуемся доказываемой в курсе математики формулой

$$\operatorname{rot} \operatorname{rot} \vec{E} = \operatorname{grad} \operatorname{div} \vec{E} - \Delta \vec{E},$$

где  $\Delta$  — введенный в главе 27 лапласиан. Первое слагаемое в правой части равно нулю в силу другого уравнения Максвелла:  $\operatorname{div} \vec{E} = 0$ . Получаем в итоге

$$-\Delta \vec{E} = -\frac{\partial}{\partial t} \operatorname{rot} \vec{B}. \quad (32.24)$$

Выразим  $\operatorname{rot} \vec{B}$  через электрическое поле с помощью уравнения Максвелла

$$\operatorname{rot} \vec{B} = \mu \mu_0 \operatorname{rot} \vec{H} = \mu \mu_0 \frac{\partial \vec{D}}{\partial t} = \mu \varepsilon \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \quad (32.25)$$

и используем это выражение в правой части (32.24). В результате приходим к уравнению

$$-\Delta \vec{E} = -\varepsilon\varepsilon_0\mu\mu_0 \frac{\partial^2 \vec{E}}{\partial t^2}. \quad (32.26)$$

Учитывая связь  $\varepsilon_0\mu_0 = 1/c^2$  и вводя **показатель преломления** среды  $n = \sqrt{\varepsilon\mu}$ , записываем уравнение для вектора напряженности электрического поля в виде

$$\frac{\partial^2 \vec{E}}{\partial x^2} + \frac{\partial^2 \vec{E}}{\partial y^2} + \frac{\partial^2 \vec{E}}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 \vec{E}}{\partial t^2}. \quad (32.27)$$

Сравнивая с (27.69), убеждаемся, что мы получили волновое уравнение, где  $v$  — **фазовая скорость света в среде**:

$$v = \frac{c}{n} = \frac{c}{\sqrt{\varepsilon\mu}}. \quad (32.28)$$

Взяв ротор от обеих частей уравнения Максвелла  $\text{rot } \vec{H} = \partial \vec{D} / \partial t$  и действуя так же, как ранее, придем к волновому уравнению для магнитного поля:

$$\frac{\partial^2 \vec{H}}{\partial x^2} + \frac{\partial^2 \vec{H}}{\partial y^2} + \frac{\partial^2 \vec{H}}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 \vec{H}}{\partial t^2}. \quad (32.29)$$

Полученные волновые уравнения для  $\vec{E}$  и  $\vec{H}$  означают, что электромагнитные поля могут существовать в виде электромагнитных волн, фазовая скорость которых равна  $v = c/n = c/\sqrt{\varepsilon\mu}$ . В отсутствие среды (при  $\varepsilon = \mu = 1$ ) скорость электромагнитных волн совпадает со скоростью света в вакууме.

## 32.4 Основные свойства электромагнитных волн

Рассмотрим плоскую монохроматическую электромагнитную волну, распространяющуюся вдоль оси  $x$ :

$$\begin{aligned} \vec{E} &= \vec{E}_0 \cos(\omega t - kx + \alpha_1), \\ \vec{H} &= \vec{H}_0 \cos(\omega t - kx + \alpha_2). \end{aligned} \quad (32.30)$$

Возможность существования таких решений следует из полученных волновых уравнений. Однако напряженности электрического и магнитного полей не являются независимыми друг от друга. Связь между ними можно установить, подставляя решения (32.30) в уравнения Максвелла.

Дифференциальную операцию  $\text{rot}$ , применяемую к некоторому векторному полю  $\vec{A}$ , можно символически записать как детерминант:

$$\begin{aligned} \text{rot } \vec{A} &= \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} = \\ &= \vec{e}_x \left( \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \vec{e}_y \left( \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \vec{e}_z \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right). \end{aligned} \quad (32.31)$$

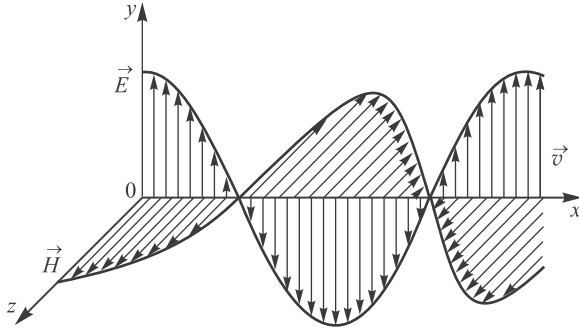


Рис. 32.1. Колебания полей  $\vec{E}$  и  $\vec{H}$  в плоской электромагнитной волне

Подставляя сюда выражения (32.30), зависящие только от  $x$ , находим:

$$\begin{aligned} \operatorname{rot} \vec{E} &= -\vec{e}_y \frac{\partial E_z}{\partial x} + \vec{e}_z \frac{\partial E_y}{\partial x} = (-\vec{e}_y E_{0z} + \vec{e}_z E_{0y}) k \sin(\omega t - kx + \alpha_1), \\ \operatorname{rot} \vec{H} &= -\vec{e}_y \frac{\partial H_z}{\partial x} + \vec{e}_z \frac{\partial H_y}{\partial x} = (-\vec{e}_y H_{0z} + \vec{e}_z H_{0y}) k \sin(\omega t - kx + \alpha_2). \end{aligned} \quad (32.32)$$

Дифференцирование плоских волн по времени дает:

$$\begin{aligned} \frac{\partial \vec{E}}{\partial t} &= -\vec{E}_0 \omega \sin(\omega t - kx + \alpha_1), \\ \frac{\partial \vec{H}}{\partial t} &= -\vec{H}_0 \omega \sin(\omega t - kx + \alpha_2). \end{aligned} \quad (32.33)$$

Тогда из уравнений Максвелла получаем

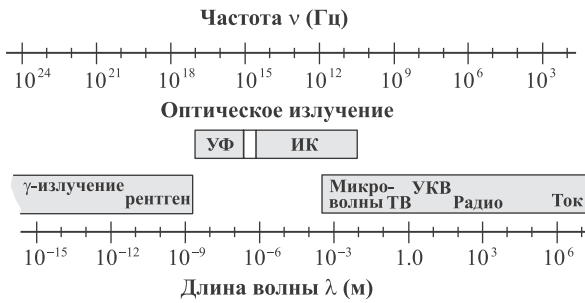
$$\begin{aligned} (-\vec{e}_y E_{0z} + \vec{e}_z E_{0y}) k \sin(\omega t - kx + \alpha_1) &= \mu \mu_0 \vec{H}_0 \omega \sin(\omega t - kx + \alpha_2), \\ (-\vec{e}_y H_{0z} + \vec{e}_z H_{0y}) k \sin(\omega t - kx + \alpha_2) &= -\varepsilon \varepsilon_0 \vec{E}_0 \omega \sin(\omega t - kx + \alpha_1). \end{aligned} \quad (32.34)$$

Отсюда следует, во-первых, что электрическое и магнитное поля колеблются в фазе:  $\alpha_1 = \alpha_2$ . Далее, ни у  $\vec{E}_0$ , ни у  $\vec{H}_0$  нет компонентов вдоль оси  $x$ :  $E_{0x} = H_{0x} = 0$ . Иными словами, **электромагнитные волны поперечны: колебания векторов электрического и магнитного полей происходят в плоскости, ортогональной направлению распространения волны.** Тогда можно выбрать координатные оси так, чтобы вектор  $\vec{E}_0$  был направлен вдоль оси  $y$ :  $\vec{E}_0 = \vec{e}_y E_0$ ,  $E_{0x} = E_{0z} = 0$  (рис. 32.1).

В этом случае уравнения (32.34) приобретают вид

$$\begin{aligned} \vec{e}_z E_0 k &= \mu \mu_0 \vec{H}_0 \omega, \\ (-\vec{e}_y H_{0z} + \vec{e}_z H_{0y}) k &= -\varepsilon \varepsilon_0 \vec{e}_y E_0 \omega. \end{aligned} \quad (32.35)$$

Отсюда следует, что вектор  $\vec{H}_0$  направлен вдоль оси  $z$ :  $\vec{H}_0 = \vec{e}_z H_0$ ,  $H_{0x} = H_{0y} = 0$ . Иначе, векторы электрического и магнитного полей ортогональны друг другу и



**Рис. 32.2.** Спектр электромагнитных волн. Диапазон видимого света выделен белым в оптической области спектра. Слева от него лежит область ультрафиолетового (УФ) излучения, справа — инфракрасного (ИК)

оба — направлению распространения волны. С учетом этого факта уравнения (32.35) еще более упрощаются:

$$\begin{aligned} E_0 k &= \mu\mu_0 H_0 \omega, \\ -H_0 k &= -\varepsilon\varepsilon_0 E_0 \omega. \end{aligned} \quad (32.36)$$

Из (32.36) вытекает обычная связь волнового вектора, частоты и скорости:

$$\frac{\omega}{k} = v = \frac{1}{\sqrt{\varepsilon_0\mu_0\varepsilon\mu}} = \frac{c}{n}. \quad (32.37)$$

Уравнения (32.36) дают нам также связь амплитуд колебаний полей:

$$\boxed{\sqrt{\varepsilon_0\varepsilon}E_0 = \sqrt{\mu_0\mu}H_0.} \quad (32.38)$$

Итак, из уравнений Максвелла следует, что электромагнитные волны распространяются в вакууме со скоростью света. В свое время этот вывод произвел огромное впечатление. Стало ясно, что не только электричество и магнетизм являются разными проявлениями одного и того же взаимодействия. Все световые явления, оптика, также стали предметом теории электромагнетизма. Различия в восприятии человеком электромагнитных волн связаны с их частотой или длиной волны. На рисунке 32.2 показаны длины волн  $\lambda$  и частоты  $\nu = c/\lambda$  характерных видов электромагнитного излучения в вакууме.

Обсуждая обнаруженное в экспериментах явление поляризации света, физики начала XIX в. говорили о поперечных колебаниях гипотетического эфира как об эмпирическом факте. Теперь же фундаментальные уравнения Максвелла привели нас к выводу о поперечности электромагнитных волн: векторы электрического и магнитного полей ортогональны друг другу и направлению распространения волны (составляют правую тройку векторов).

### 32.5 Эффект Доплера для электромагнитных волн

Пусть в некоторой инерциальной системе отсчета  $K$  распространяется плоская электромагнитная волна. Фаза волны имеет вид

$$\varphi = \omega t - \vec{k} \cdot \vec{r} + \varphi_0. \quad (32.39)$$

Наблюдатель в другой инерциальной системе отсчета  $K'$ , движущейся относительно первой со скоростью  $V$  вдоль оси  $x$ , также наблюдает эту волну, но пользуется другими координатами и временем:  $t'$ ,  $\vec{r}'$ . Связь между системами отсчета дается преобразованиями Лоренца (см. т. 1, разд. 9.3):

$$x = \frac{x' + Vt'}{\sqrt{1 - V^2/c^2}}, \quad y = y', \quad z = z', \quad t = \frac{t' + (V/c^2)x'}{\sqrt{1 - V^2/c^2}}. \quad (32.40)$$

При переходе к другой системе отсчета фаза волны не должна меняться: очевидно, что гребень останется гребнем, а впадина — впадиной. Подставим выражения (32.40) в (32.39), чтобы получить фазу волны в движущейся системе отсчета:

$$\begin{aligned} \varphi &= \omega \frac{t' + (V/c^2)x'}{\sqrt{1 - V^2/c^2}} - k_x \frac{x' + Vt'}{\sqrt{1 - V^2/c^2}} - k_y y' - k_z z' + \varphi_0 = \\ &= \frac{\omega - k_x V}{\sqrt{1 - V^2/c^2}} t' - \frac{k_x - \omega V/c^2}{\sqrt{1 - V^2/c^2}} x' - k_y y' - k_z z' + \varphi_0. \end{aligned} \quad (32.41)$$

Это выражение можно записать как

$$\varphi = \omega' t' - \vec{k}' \cdot \vec{r}' + \varphi_0, \quad (32.42)$$

где  $\omega'$  и  $\vec{k}'$  — циклическая частота и волновой вектор относительно движущейся системы отсчета. Сравнивая с (32.41), находим преобразования Лоренца для частоты и волнового вектора:

$$\omega' = \frac{\omega - k_x V}{\sqrt{1 - V^2/c^2}}, \quad k'_x = \frac{k_x - \omega V/c^2}{\sqrt{1 - V^2/c^2}}, \quad k'_y = k_y, \quad k'_z = k_z. \quad (32.43)$$

Для электромагнитной волны в вакууме  $\omega/k = c$ . Пусть направление распространения волны составляет в первой системе отсчета угол  $\alpha$  с осью  $x$ :  $k_x = k \cos \alpha = (\omega/c) \cos \alpha$ . Тогда выражение для частоты волны в движущейся системе отсчета принимает вид

$$\omega' = \frac{\omega}{\sqrt{1 - V^2/c^2}} \left( 1 - \frac{V}{c} \cos \alpha \right). \quad (32.44)$$

Если  $\alpha = 0$ , то наблюдатель удаляется от источника излучения и воспринимаемая им частота волны уменьшается:

$$\omega' = \omega \frac{1 - V/c}{\sqrt{1 - V^2/c^2}}. \quad (32.45)$$

Если  $\alpha = \pi$ , то наблюдатель приближается к источнику и частота излучения для него увеличивается:

$$\omega' = \omega \frac{1 + V/c}{\sqrt{1 - V^2/c^2}}. \quad (32.46)$$

При скоростях  $V \ll c$  можно пренебречь отклонением квадратного корня в знаменателях от единицы, и мы приходим к формуле, аналогичной (27.85) для эффекта Доплера в звуковой волне.

Отметим существенную особенность эффекта Доплера для электромагнитной волны. Скорость движущейся системы отсчета играет здесь роль относительной скорости наблюдателя и источника. Полученные формулы автоматически удовлетворяют принципу относительности Галилея, и с помощью экспериментов невозможно установить, что именно движется — источник или наблюдатель. Это связано с тем, что для электромагнитных волн отсутствует среда (эфир), которая играла бы ту же роль, что и воздух для звуковой волны. Заметим также, что для электромагнитных волн имеет место *поперечный эффект Доплера*. При  $\alpha = \pi/2$  частота излучения изменяется:

$$\omega' = \frac{\omega}{\sqrt{1 - V^2/c^2}}, \quad (32.47)$$

в то время как для звуковых волн движение в направлении, ортогональном распространению волны, не приводит к сдвигу частот. Этот эффект прямо связан с релятивистским замедлением времени в движущейся системе отсчета: наблюдатель на ракете видит увеличение частоты излучения или, более общо, ускорение всех процессов, происходящих на Земле.

Найдем теперь фазовую скорость волны  $c' = \omega'/k'$  в движущейся системе отсчета. Имеем из преобразований Лоренца для волнового вектора

$$k'^2 = k_x'^2 + k_y'^2 + k_z'^2 = \frac{(k_x - \omega V/c^2)^2}{1 - V^2/c^2} + k_y^2 + k_z^2. \quad (32.48)$$

Подставим сюда соотношение

$$k_y^2 + k_z^2 = k^2 - k_x^2 = \frac{\omega^2}{c^2} - k_x^2. \quad (32.49)$$

Получаем

$$\begin{aligned} k'^2 &= \frac{(k_x - \omega V/c^2)^2}{1 - V^2/c^2} + \frac{\omega^2}{c^2} - k_x^2 = \\ &= \frac{(k_x - \omega V/c^2)^2 + (\omega^2/c^2)(1 - V^2/c^2) - k_x^2(1 - V^2/c^2)}{1 - V^2/c^2} = \\ &= \frac{(\omega - k_x V)^2}{c^2(1 - V^2/c^2)} = \frac{\omega'^2}{c^2}. \end{aligned} \quad (32.50)$$

Отсюда находим скорость волны в движущейся системе отсчета:

$$c'^2 = \frac{\omega'^2}{k'^2} = c^2. \quad (32.51)$$

Мы обнаружили, что скорость волны в движущейся системе отсчета не изменилась и по-прежнему равна скорости света  $c$ , что, как и должно быть, соответствует постулату специальной теории относительности.

**Задача 32.28.** Фотонная ракета движется со скоростью  $V = 0,9c$ , держа курс на звезду, наблюдающуюся с Земли в оптическом диапазоне (длина волны  $\lambda \sim 0,4$  мкм). Какой увидят звезду космонавты?

**Решение.** Длина волны обратно пропорциональна частоте колебаний. Из формулы (32.46) для эффекта Доплера в случае сближения источника света и наблюдателя находим закон преобразования длин волн:

$$\lambda' = \lambda \sqrt{\frac{1 - V/c}{1 + V/c}}. \quad (32.52)$$

Отсюда следует результат:

$$\lambda' = 0,4 \sqrt{\frac{1 - 0,9}{1 + 0,9}} = 0,09 \text{ мкм}. \quad (32.53)$$

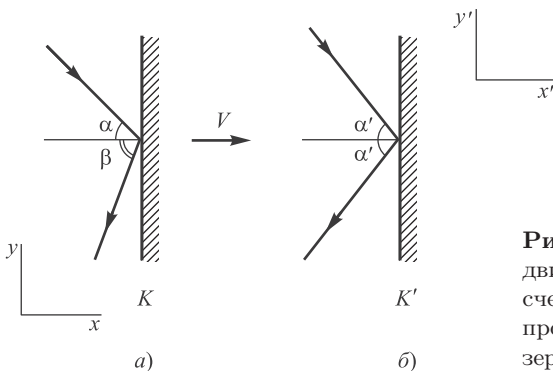
По рисунку 32.2 определяем, что для космонавтов излучение звезды сместилось в ультрафиолетовый диапазон. ■

## Отражение света от движущегося зеркала

В первом томе учебника (см. разд. 4.7) мы проанализировали упругое отражение шара от движущейся стенки. Здесь мы рассмотрим релятивистский аналог этого процесса — отражение света от движущегося зеркала. Пусть зеркало движется в вакууме перпендикулярно своей плоскости с постоянной скоростью  $V$ . На него под углом  $\alpha$  к нормали падает световая волна частотой  $\omega$ . Найдем, какой угол с нормалью составит отраженный луч и как изменится при отражении частота света.

Сначала выберем координатные оси, как показано на рис. 32.3, *a*. Тогда компоненты волнового вектора можно записать как

$$\vec{k} = \frac{\omega}{c} \{ \cos \alpha, -\sin \alpha, 0 \}. \quad (32.54)$$



**Рис. 32.3.** Отражение света от зеркала, движущегося в лабораторной системе отсчета  $K$  со скоростью  $V$  (*a*), и тот же процесс в собственной системе отсчета зеркала (*b*)

Волновой вектор  $\vec{k}'$  и частота  $\omega'$  в системе отсчета  $K'$ , где зеркало покоится, даются преобразованиями Лоренца (32.43):

$$\omega' = \omega \frac{1 - (V/c) \cos \alpha}{\sqrt{1 - V^2/c^2}}, \quad \vec{k}' = \frac{\omega}{c} \left\{ \frac{\cos \alpha - V/c}{\sqrt{1 - V^2/c^2}}, -\sin \alpha, 0 \right\}. \quad (32.55)$$

Волновой вектор отраженного света обозначим  $\vec{q}'$ , а частоту —  $\omega'_f$ . При отражении от покоящегося (в системе  $K'$ ) зеркала частота не меняется ( $\omega'_f = \omega'$ ), а угол падения  $\alpha'$  равен углу отражения (рис. 32.3, б). Это значит, что тангенциальные компоненты волнового вектора не меняются, а его нормальная компонента изменяет знак, т. е. в системе  $K'$  имеем:

$$\omega'_f = \omega \frac{1 - (V/c) \cos \alpha}{\sqrt{1 - V^2/c^2}}, \quad \vec{q}' = \frac{\omega}{c} \left\{ -\frac{\cos \alpha - V/c}{\sqrt{1 - V^2/c^2}}, -\sin \alpha, 0 \right\}. \quad (32.56)$$

Чтобы вернуться назад в лабораторную систему отсчета, надо применить преобразования Лоренца, обратные преобразованиям (32.43):

$$\omega_f = \frac{\omega'_f + q'_x V}{\sqrt{1 - V^2/c^2}}, \quad q_x = \frac{q'_x + \omega'_f V/c^2}{\sqrt{1 - V^2/c^2}}, \quad q_y = q'_y, \quad q_z = q'_z. \quad (32.57)$$

Подставляя сюда выражения (32.56), находим частоту отраженного света

$$\omega_f = \omega \frac{1 - 2(V/c) \cos \alpha + V^2/c^2}{1 - V^2/c^2} \quad (32.58)$$

и компоненты волнового вектора:

$$q_x = \frac{\omega}{c} \frac{2V/c - \cos \alpha (1 + V^2/c^2)}{1 - V^2/c^2}, \quad q_y = -\frac{\omega}{c} \sin \alpha, \quad q_z = 0. \quad (32.59)$$

Тангенс угла отражения можно найти как отношение компонентов волнового вектора  $q_y/q_x$ , откуда

$$\operatorname{tg} \beta = \frac{(1 - V^2/c^2) \sin \alpha}{(1 + V^2/c^2) \cos \alpha - 2V/c}. \quad (32.60)$$

Если пренебречь членами второго порядка  $V^2/c^2$ , то приходим к формуле (4.80) из т. 1.

В частном случае нормального падения луча света на зеркало формулы (32.58)–(32.60) принимают вид

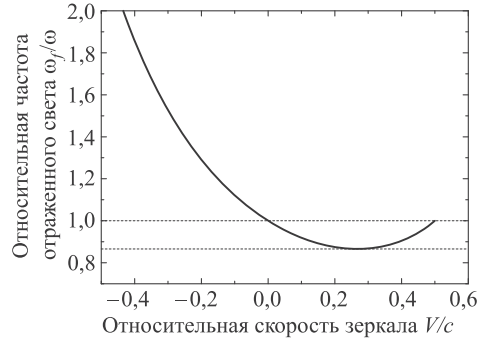
$$\omega_f = \omega \frac{1 - V/c}{1 + V/c}, \quad \vec{q} = \frac{\omega_f}{c} \{-1, 0, 0\}, \quad \beta = 0. \quad (32.61)$$

Если зеркало движется навстречу световому лучу ( $V < 0$ ), частота отраженного луча возрастает и стремится к бесконечности по мере приближения скорости зеркала к скорости света. Если зеркало «убегает» от светового луча, то частота





**Рис. 32.4.** Зависимость угла отражения от скорости зеркала. Пунктирные горизонтальные линии отмечают характерные значения:  $\beta = \alpha$  при  $V = 0$ ,  $\beta = \pi/2$  при  $V = c(1 - \sin \alpha)/\cos \alpha$  и  $\beta = \pi - \alpha$  при  $V = c \cos \alpha$



**Рис. 32.5.** Зависимость частоты отраженного света от скорости зеркала. Пунктирные горизонтальные линии отмечают характерные значения:  $\omega_f = \omega$  при  $V = 0$  и  $V = c \cos \alpha$ , а также  $\omega_f = \omega \sin \alpha$  при  $\beta = \pi/2$ ,  $V = c(1 - \sin \alpha)/\cos \alpha$

отраженного света уменьшается и стремится к нулю при  $V \rightarrow c$ . Можно считать, что отраженный луч испускается мнимым источником, являющимся отражением в зеркале реального источника. С этой точки зрения мы имеем дело с эффектом Доплера: с «фиолетовым смещением» частоты при приближении источника света к наблюдателю или с «красным смещением» при его удалении.

При скорости зеркала  $V = c \cos \alpha$  луч уже не может «догнать» зеркало, и из наших формул следует, что  $\omega_f = \omega$ ,  $\vec{q} = \vec{k}$  и  $\beta = \pi - \alpha$ , т. е. луч просто не меняет своего направления.

Если скорость зеркала удовлетворяет квадратному уравнению

$$2V/c - \cos \alpha(1 + V^2/c^2) = 0, \text{ т. е. } V = c \frac{1 - \sin \alpha}{\cos \alpha}, \quad (32.62)$$

то  $q_x = 0$  и отраженный луч движется параллельно плоскости зеркала ( $\beta = \pi/2$ ). Частота отраженного света в этом случае равна  $\omega_f = \omega \sin \alpha$ . В промежутке скоростей зеркала

$$c \frac{1 - \sin \alpha}{\cos \alpha} < V < c \cos \alpha$$

угол отражения заключен в интервале  $\pi/2 < \beta < \pi - \alpha$ ; луч словно преломляется, следуя за зеркалом. На рисунках 32.4 и 32.5 показаны соответственно зависимости угла отражения и частоты отраженного луча от скорости зеркала для конкретного значения угла падения  $\alpha = \pi/3$ .

## 32.6 Энергия и импульс электромагнитного поля

Объемная плотность энергии  $w$  электромагнитной волны складывается из объемных плотностей энергии  $w_{\text{эл}}$  электрического и  $w_{\text{м}}$  магнитного полей:

$$w = w_{\text{эл}} + w_{\text{м}} = \frac{\varepsilon_0 \varepsilon E^2}{2} + \frac{\mu_0 \mu H^2}{2}. \quad (32.63)$$

Учитывая связь векторов  $\vec{E}$  и  $\vec{H}$ , получим, что плотности энергии электрического и магнитного полей в каждый момент времени одинаковы, т. е.  $w_{\text{эл}} = w_{\text{м}}$ . Следовательно,  $w$  можно представить в виде

$$w = 2w_{\text{эл}} = \sqrt{\varepsilon_0\mu_0}\sqrt{\varepsilon\mu}EH. \quad (32.64)$$

Если умножить плотность энергии  $w$  на скорость электромагнитной волны в среде  $v = c/\sqrt{\varepsilon\mu}$ , то получим модуль плотности потока энергии:

$$S = wv = EH. \quad (32.65)$$

Так как векторы  $\vec{E}$  и  $\vec{H}$  взаимно перпендикулярны и образуют с направлением распространения волны правовинтовую систему, направление вектора  $\vec{E} \times \vec{H}$  совпадает с направлением распространения волны, т. е. с направлением переноса энергии, а модуль этого вектора равен  $EH$ . Следовательно, вектор плотности потока электромагнитной энергии (**вектор Пойнтинга**<sup>1</sup>) имеет вид

$$\vec{S} = \vec{v} \cdot w = \vec{E} \times \vec{H}. \quad (32.66)$$

Как и для звуковых волн, интенсивность электромагнитной волны — это среднее значение плотности потока энергии:

$$\begin{aligned} I &= \langle S \rangle = \langle EH \rangle = E_0H_0 \langle \cos^2(\omega t - kx + \varphi_0) \rangle = \frac{1}{2}E_0H_0 = \\ &= \frac{c}{\sqrt{\varepsilon\mu}} \cdot \frac{\varepsilon_0\varepsilon E_0^2}{2} = \frac{c}{\sqrt{\varepsilon\mu}} \cdot \frac{\mu_0\mu H_0^2}{2}. \end{aligned} \quad (32.67)$$

Как и в звуковой волне, интенсивность пропорциональна квадрату амплитуды колебаний.

**Задача 32.29.** Интенсивность солнечного излучения, падающего на Землю, составляет  $I = 1,4$  кВт/м<sup>2</sup> (так называемая солнечная постоянная). Найти среднюю амплитуду колебаний  $E_0$  вектора электрической напряженности в солнечном излучении. Вычислить амплитуды колебаний напряженности магнитного поля  $H_0$  и вектора магнитной индукции  $B_0$  в волне.

**Решение.** Ответ находим сразу из уравнений (32.67), где полагаем  $\varepsilon = \mu = 1$ :

$$E_0 = \sqrt{\frac{2I}{c\varepsilon_0}} = \sqrt{\frac{2 \times 1,4 \cdot 10^3}{3 \cdot 10^8 \times 8,85 \cdot 10^{-12}}} = 1027 \text{ В/м} \approx 1 \text{ кВ/м},$$

$$H_0 = \sqrt{\frac{2I}{c\mu_0}} = \sqrt{\frac{2 \times 1,4 \cdot 10^3}{3 \cdot 10^8 \times 4\pi \cdot 10^{-7}}} = 2,7 \text{ А/м},$$

$$B_0 = \mu_0 H_0 = 2,7 \times 4\pi \cdot 10^{-7} = 3,4 \cdot 10^{-6} = 3,4 \text{ мкТл}. \quad \blacksquare$$

<sup>1</sup>По российской научной традиции называется вектором Умова—Пойнтинга. Британский физик Д. Пойнтинг ввел этот вектор для электромагнитного поля (1884), но Н. А. Умов считается предтечей Пойнтинга, так как еще в 1874 г. в своей докторской диссертации использовал новое для того времени понятие плотности энергии и нашел уравнения распространения энергии в упругих средах. О колоссальном впечатлении, которое Умов производил на студентов, можно судить по стихам поэта Андрея Белого (см. эпиграф к части II, т. 1).

### Давление света

Электромагнитные волны поглощаются или отражаются телами, следовательно, они должны оказывать на тела давление. Рассмотрим плоскую электромагнитную волну, падающую нормально на плоскую проводящую поверхность. В этом случае электрическое поле волны возбуждает в проводнике ток, пропорциональный  $\vec{E}$ . Магнитное поле волны по закону Ампера будет действовать на ток с силой, направление которой совпадает с направлением распространения волны. В 1899 г. в исключительно тонких экспериментах П. Н. Лебедев доказал существование светового давления. Можно показать, что волна, несущая энергию  $W$ , обладает и импульсом

$$p = \frac{W}{c}. \quad (32.68)$$

Аналогичная связь между импульсом и полной энергией существует для частиц с нулевой массой (см. т. 1, разд. 9.12), к которым относятся кванты света — фотоны.

При падении по нормали на какую-то поверхность часть энергии волны  $W_{\text{отр}} = \alpha W$  отразится в противоположном направлении, а часть  $W_{\text{погл}} = \eta W$  поглотится веществом поверхности. Параметры  $\alpha$  и  $\eta$  называются *коэффициентами отражения* и *поглощения* соответственно. Поглощенная часть волны передает поверхности весь свой импульс, а отраженная меняет свой импульс на противоположный и, стало быть, передает поверхности удвоенный импульс.

Пусть электромагнитная волна падает в вакууме по нормали на поверхность площадью  $A$ . Предположим, что за время  $\Delta t$  волна переносит энергию  $\Delta W$ . Тогда переданный площадке импульс равен:

$$\Delta p = \frac{\Delta W}{c}(2\alpha + \eta).$$

По второму закону Ньютона на площадку действует со стороны волны сила

$$F = \frac{\Delta p}{\Delta t} = \frac{\Delta W}{c\Delta t}(2\alpha + \eta).$$

Тогда давление  $P$  волны на поверхность равно:

$$P = \frac{F}{A} = \frac{\Delta W}{Ac\Delta t}(2\alpha + \eta).$$

Если средняя плотность энергии в волне равна  $\langle w \rangle$ , то на площадку  $A$  за время  $\Delta t$  попадет энергия из объема  $Ac\Delta t$  и  $\Delta W = (Ac\Delta t)\langle w \rangle$ . Отсюда находим связь давления  $P$  электромагнитной волны с ее интенсивностью  $I$ :

$$P = \langle w \rangle(2\alpha + \eta) = \frac{I}{c}(2\alpha + \eta). \quad (32.69)$$

Если поверхность поглощает всю падающую на нее волну (абсолютно черное тело), то  $\alpha = 0$ ,  $\eta = 1$  и давление волны равно  $P = I/c$ ; в обратном случае

идеально отражающей поверхности  $\alpha = 1$ ,  $\eta = 0$ , и тогда давление будет в 2 раза больше:  $P = 2I/c$ .

Заметим, что помимо отражения и поглощения падающая волна может проходить сквозь прозрачное тело:  $W_{\text{прох}} = dW$ , где  $d$  — коэффициент прохождения. Давления эта часть волны не оказывает. А по закону сохранения энергии сумма всех трех коэффициентов должна быть равна единице:  $\alpha + \eta + d = 1$ .

**Задача 32.30.** Найти давление  $P$  солнечного света на часть поверхности Земли, ортогональную солнечным лучам. Коэффициент отражения Земли  $\alpha = 0,3$ .

**Решение.** Так как Земля непрозрачна, то коэффициент прохождения  $d = 0$ , а коэффициент поглощения  $\eta = 1 - \alpha = 0,7$ . Используем значение солнечной постоянной из предыдущей задачи. Искомое давление равно:

$$P = \frac{I}{c}(2\alpha + \eta) = \frac{1,4 \cdot 10^3}{3 \cdot 10^8}(2 \times 0,3 + 0,7) = 6,1 \cdot 10^{-6} \text{ Па} = 6,1 \text{ мкПа}. \blacksquare$$

**Задача 32.31.** Найти давление  $P$  лазерного пучка на поглощающую мишень. Выходная мощность лазера  $N = 4,6$  Вт, диаметр пучка  $d = 2,6$  мм.

**Решение.** Площадь сечения пучка лазерного излучения  $A = \pi d^2/4$ , интенсивность излучения  $I = N/A = 4N/(\pi d^2)$ . Отсюда находим:

$$P = \frac{I}{c} = \frac{4N}{\pi d^2 c} = \frac{4 \times 4,6}{\pi \times (2,6 \cdot 10^{-3})^2 \times 3 \cdot 10^8} = 2,9 \cdot 10^{-3} \text{ Па} = 2,9 \text{ мПа}. \blacksquare$$

## Солнечный парус

Посмотрим теперь, какие изменения надо внести в полученные формулы, если излучение падает на площадку  $A$  под углом  $\vartheta$  к нормали. Прежде всего надо учесть, что площадь поперечного сечения пучка равна теперь  $A \cos \vartheta$ , так что за время  $\Delta t$  на площадку попадет энергия

$$\Delta W = (Ac\Delta t)\langle w \rangle \cos \vartheta.$$

Переносимый волной импульс имеет нормальную составляющую

$$p_n = \frac{\Delta W}{c} \cos \vartheta = (A\Delta t)\langle w \rangle \cos^2 \vartheta$$

и тангенциальную составляющую

$$p_\tau = \frac{\Delta W}{c} \sin \vartheta = (A\Delta t)\langle w \rangle \sin \vartheta \cos \vartheta.$$

От нормальной составляющей площадки передается импульс  $2\alpha p_n$  при отражении части излучения и импульс  $\eta p_n$  при прохождении другой части. В итоге в направлении нормали за время  $\Delta t$  площадке  $A$  передается импульс

$$\Delta p_n = p_n(2\alpha + \eta) = (A\Delta t)\langle w \rangle(2\alpha + \eta) \cos^2 \vartheta$$

и на площадку действует нормальная сила  $F_n = \Delta p_n/\Delta t$ . Таким образом,

$$F_n = A\langle w \rangle(2\alpha + \eta) \cos^2 \vartheta = A \frac{I}{c}(2\alpha + \eta) \cos^2 \vartheta. \quad (32.70)$$

При отражении тангенциальная составляющая импульса не меняется и площадке ничего не передается. Но из-за поглощения излучения площадке передается тангенциальный импульс

$$\Delta p_\tau = \eta p_\tau = (A\Delta t)\langle w \rangle \eta \sin \vartheta \cos \vartheta.$$

Следовательно, на площадку действует тангенциальная сила  $F_\tau = \Delta p_\tau / \Delta t$ , равная

$$F_\tau = A \frac{I}{c} \eta \sin \vartheta \cos \vartheta. \quad (32.71)$$

Отношение тангенциальной и нормальной составляющих определяет угол  $\theta$  между нормалью к площадке и направлением силы со стороны падающего излучения:

$$\operatorname{tg} \theta = \frac{F_\tau}{F_n} = \frac{\eta}{2\alpha + \eta} \operatorname{tg} \vartheta. \quad (32.72)$$

Если площадка полностью поглощает излучение  $\eta = 1, \alpha = 0$ , то сила направлена вдоль падающего излучения ( $\theta = \vartheta$ ). Если же, наоборот, площадка сделана из идеально отражающего материала ( $\alpha = 1, \eta = 0$ ), то сила действует по нормали к площадке ( $\theta = 0$ ).

Сила давления на площадку со стороны солнечного света подобна силам, действующим на парус яхты. Идею использования *солнечного паруса* для полетов в космосе высказал в 1924 г. советский ученый Ф. А. Цандер, один из пионеров ракетостроения.

В 2005 г. в рамках совместного российско-американского проекта с подводной лодки был произведен пуск ракеты со спутником «Космос-1». Спутник был оснащен солнечным парусом, состоящим из восьми лепестков общим диаметром 30 м. Целью эксперимента была проверка возможности использования солнечного паруса в качестве движителя аппарата. Однако запуск оказался неудачным и аппарат разбился в океане.

В 2010 г. Японское агентство аэрокосмических исследований запустило космический аппарат *IKAROS*, ставший первым космическим аппаратом, оснащенным солнечным парусом. Размер паруса  $14 \times 14 \text{ м}^2$ , он изготовлен из полиамидной пленки толщиной 7,5 мкм. Система разворачивания паруса сработала в штатном режиме, и через полтора месяца после запуска ученые сообщили, что солнечный свет давит на *IKAROS* с силой 1,12 мН. Под действием давления света аппарат массой 315 кг приобретает ускорение  $a = 3,6 \cdot 10^{-6} \text{ м/с}^2$ . При длительном полете даже такое малое ускорение может привести к существенному увеличению скорости без расходования топлива. В планах японских ученых — создание солнечного паруса с площадью поверхности в 10 раз большей, чем у *IKAROS*'а. Они полагают, что с его помощью можно без расходования топлива увеличить скорость космического аппарата на несколько километров в секунду, что сравнимо с результатами гравитационного маневра (см. т. 1, разд. 6.7).

## 32.7 Законы оптики и уравнения Максвелла.

### Формулы Френеля

Волновая теория света, берущая свое начало от Х. Гюйгенса, основана на предположении, что свет представляет собой волновой процесс. Какая среда при этом колеблется, во времена Гюйгенса было непонятно. Эту среду — мировой эфир — представляли себе механистически, а его колебания — световые волны — подобными колебаниям твердого тела. Исходя из таких самых общих представлений о волновых процессах удалось описать явления интерференции, дифракции и поляризации. Волновая теория света сначала была подтверждена в многочисленных экспериментах и только позже нашла полное теоретическое описание в рамках классической электродинамики Максвелла. В этом разделе мы покажем, как из общих уравнений Максвелла с граничными условиями на поверхности соприкосновения различных сред и из свойств электромагнитных волн вытекают некоторые законы оптики.

Поскольку электромагнитную волну можно представить себе как суперпозицию двух линейно-поляризованных волн, мы рассмотрим отдельно два случая.

## Отражение и преломление $s$ -поляризованной волны

Рассмотрим плоскую электромагнитную волну, падающую на плоскую границу раздела двух однородных и изотропных сред 1 и 2, характеризующихся разными значениями диэлектрической и магнитной проницаемостей  $\epsilon_1$  и  $\epsilon_2$  (для простоты будем считать среды немагнитными диэлектриками:  $\mu_1 = \mu_2 = 1$ ). Предположим сначала, что эта волна поляризована ортогонально плоскости падения (рис. 32.6). Такая волна называется  $s$ -поляризованной.

Возьмем какую-нибудь точку на поверхности раздела двух сред. Фазы всех трех волн в ней одинаковы, и достаточно рассмотреть их амплитуды. Из геометрии следует связь с углами падения  $i_{\text{п}}$ , отражения  $i_{\text{от}}$  и преломления  $i_{\text{пр}}$  проекций амплитуд напряженностей магнитного поля падающей  $H_1$ , отраженной  $H'_1$  и преломленной  $H_2$  волн на координатные оси:

$$\begin{aligned} H_{1y} &= -H_1 \sin i_{\text{п}}, & H_{1x} &= -H_1 \cos i_{\text{п}}; \\ H'_{1y} &= -H'_1 \sin i_{\text{от}}, & H'_{1x} &= H'_1 \cos i_{\text{от}}; \\ H_{2y} &= -H_2 \sin i_{\text{пр}}, & H_{2x} &= -H_2 \cos i_{\text{пр}}. \end{aligned} \quad (32.73)$$

Тангенциальные составляющие напряженности магнитного поля должны быть непрерывны на границе раздела двух сред, а нормальные, вообще говоря, терпят скачок (см. т. 1, разд. 21.5), но для немагнитных сред тоже непрерывны:

$$\begin{aligned} H_{1x} + H'_{1x} &= H_{2x}, \\ H_{1y} + H'_{1y} &= H_{2y}. \end{aligned} \quad (32.74)$$

Подставляя сюда выражения (32.73), получаем два уравнения для амплитудных значений напряженностей магнитного поля в падающей, отраженной и преломленной волнах:

$$\begin{aligned} H_1 \cos i_{\text{п}} - H'_1 \cos i_{\text{от}} &= H_2 \cos i_{\text{пр}}, \\ H_1 \sin i_{\text{п}} + H'_1 \sin i_{\text{от}} &= H_2 \sin i_{\text{пр}}. \end{aligned} \quad (32.75)$$

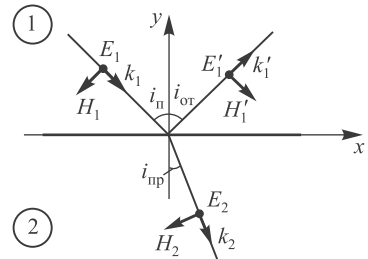
На следующем этапе мы используем связь (32.38) между напряженностями электрического и магнитного поля в волне:

$$H = \sqrt{\frac{\epsilon_0 \epsilon}{\mu_0 \mu}} E.$$

Тогда уравнения (32.75) преобразуются к виду

$$\begin{aligned} \sqrt{\epsilon_1} (E_1 \cos i_{\text{п}} - E'_1 \cos i_{\text{от}}) &= \sqrt{\epsilon_2} E_2 \cos i_{\text{пр}}, \\ \sqrt{\epsilon_1} (E_1 \sin i_{\text{п}} + E'_1 \sin i_{\text{от}}) &= \sqrt{\epsilon_2} E_2 \sin i_{\text{пр}}. \end{aligned} \quad (32.76)$$

Мы имеем два уравнения для трех амплитуд электрического поля. Третье уравнение получим из условия непрерывности тангенциальной составляющей электрического поля. Во всех трех волнах электрическое поле направлено ортогонально плоскости падения



**Рис. 32.6.** Падение  $s$ -поляризованной волны на плоскую границу раздела двух сред; векторы напряженности магнитного поля падающей, отраженной и преломленной волн лежат в плоскости падения волны (т. е. в плоскости рисунка), а соответствующие векторы напряженности электрического поля ортогональны ей, т. е. ортогональны плоскости рисунка и направлены на нас (показано точкой на луче)

волны, т. е. все три вектора параллельны между собой и параллельны поверхности раздела сред. Отсюда имеем уравнение

$$E_2 = E_1 + E'_1. \quad (32.77)$$

Подставляя это выражение для  $E_2$  в (32.76), приходим к системе уравнений для амплитуд падающей и отраженной волн:

$$\begin{aligned} E_1 (\sqrt{\varepsilon_1} \cos i_{\text{п}} - \sqrt{\varepsilon_2} \cos i_{\text{пп}}) - E'_1 (\sqrt{\varepsilon_1} \cos i_{\text{от}} + \sqrt{\varepsilon_2} \cos i_{\text{пп}}) &= 0, \\ E_1 (\sqrt{\varepsilon_1} \sin i_{\text{п}} - \sqrt{\varepsilon_2} \sin i_{\text{пп}}) + E'_1 (\sqrt{\varepsilon_1} \sin i_{\text{от}} - \sqrt{\varepsilon_2} \sin i_{\text{пп}}) &= 0. \end{aligned} \quad (32.78)$$

Эта система имеет отличные от нуля решения лишь при равенстве нулю ее детерминанта:

$$\begin{vmatrix} \sqrt{\varepsilon_1} \cos i_{\text{п}} - \sqrt{\varepsilon_2} \cos i_{\text{пп}} & -\sqrt{\varepsilon_1} \cos i_{\text{от}} - \sqrt{\varepsilon_2} \cos i_{\text{пп}} \\ \sqrt{\varepsilon_1} \sin i_{\text{п}} - \sqrt{\varepsilon_2} \sin i_{\text{пп}} & \sqrt{\varepsilon_1} \sin i_{\text{от}} - \sqrt{\varepsilon_2} \sin i_{\text{пп}} \end{vmatrix} = 0. \quad (32.79)$$

Вычитая второй столбец из первого, приводим детерминант к виду

$$\begin{vmatrix} n_1 (\cos i_{\text{п}} + \cos i_{\text{от}}) & -n_1 \cos i_{\text{от}} - n_2 \cos i_{\text{пп}} \\ n_1 (\sin i_{\text{п}} - \sin i_{\text{от}}) & n_1 \sin i_{\text{от}} - n_2 \sin i_{\text{пп}} \end{vmatrix} = 0. \quad (32.80)$$

Здесь мы ввели показатели преломления сред  $n_i = \sqrt{\varepsilon_i}$ .

Теперь очевидно, что при  $i_{\text{п}} = i_{\text{от}}$  (угол падения равен углу отражения) и  $n_1 \sin i_{\text{от}} = n_2 \sin i_{\text{пп}}$  (закон Снеллиуса) равна нулю вторая строка детерминанта и, стало быть, он сам тоже равен нулю. Мы убедились, что два закона геометрической оптики вытекают из уравнений Максвелла.

Используя эти соотношения, мы можем найти из системы (32.78) отношение амплитуд отраженной и падающей волн:

$$\frac{E'_1}{E_1} = \frac{n_1 \cos i_{\text{п}} - n_2 \cos i_{\text{пп}}}{n_1 \cos i_{\text{п}} + n_2 \cos i_{\text{пп}}}. \quad (32.81)$$

Из (32.77) теперь легко находим амплитуду преломленной волны:

$$\frac{E_2}{E_1} = \frac{2n_1 \cos i_{\text{п}}}{n_1 \cos i_{\text{п}} + n_2 \cos i_{\text{пп}}}. \quad (32.82)$$

В частности, при нормальном падении волны на поверхность раздела ( $i_{\text{п}} = i_{\text{пп}} = 0$ ) имеем

$$\frac{E'_1}{E_1} = \frac{n_1 - n_2}{n_1 + n_2} = \frac{1 - n}{1 + n}, \quad \frac{E_2}{E_1} = \frac{2n_1}{n_1 + n_2} = \frac{2}{1 + n}, \quad (32.83)$$

где  $n = n_2/n_1$  — относительный показатель преломления первой среды относительно второй. При отражении от более плотной среды ( $n > 1$ ) амплитуда отраженного света отрицательна, что означает сдвиг фазы на угол  $\pi$ . Этот результат мы учитывали при рассмотрении явления интерференции с участием отраженных волн — см. разделы 29.5 (интерференция в тонких пленках) и 29.6 (кольца Ньютона).

С помощью закона Снеллиуса можно исключить показатели преломления сред из соотношений (32.81):

$$\boxed{\begin{aligned} \frac{E'_1}{E_1} &= -\frac{\sin(i_{\text{п}} - i_{\text{пп}})}{\sin(i_{\text{п}} + i_{\text{пп}})}, \\ \frac{E_2}{E_1} &= \frac{2 \cos i_{\text{п}} \sin i_{\text{пп}}}{\sin(i_{\text{п}} + i_{\text{пп}})}. \end{aligned}} \quad (32.84)$$

Интенсивность света пропорциональна квадрату амплитуды напряженности электрического поля. Поэтому для отношения интенсивности  $I_{от}$  отраженного света к интенсивности  $I_{п}$  падающего света (эта величина  $R_s$  называется **коэффициентом отражения**) получаем выражение

$$R_s = \frac{\sin^2(i_{п} - i_{пр})}{\sin^2(i_{п} + i_{пр})}. \quad (32.85)$$

При нормальном падении луча ( $i_{п} = i_{пр} = 0$ ) из (32.81) следует, что

$$R_s = \left( \frac{n_2 - n_1}{n_2 + n_1} \right)^2 = \left( \frac{n - 1}{n + 1} \right)^2. \quad (32.86)$$

**Коэффициент прохождения** определяется из закона сохранения энергии как

$$D_s = 1 - R_s = \frac{4n}{(n + 1)^2}. \quad (32.87)$$

Законы интерференции и формулы Френеля объясняют эффект **просветления оптики**, когда на поверхность линз наносится тонкая пленка (или нескольких слоев пленок) другого материала для подавления бликов и увеличения светопропускания оптической системы. Мы рассмотрим простейший случай однослойного просветления. (Многослойные покрытия используются на практике для просветления оптики в достаточно широком диапазоне длин волн, поскольку имеет место дисперсия показателя преломления.)

Пусть свет длиной волны  $\lambda$  падает по нормали из среды с показателем преломления  $n_1$  на тонкую пленку толщиной  $d$  и показателем преломления  $n_2$  ( $n_2 > n_1$ ), за которой находится среда (линза) с показателем преломления  $n_3$  ( $n_3 > n_1$ ). Сформулируем условия, при которых пленка окажет просветляющее действие. Свет отражается от передней и задней поверхностей пленки. Чтобы при интерференции эти волны ослабляли друг друга, их фазы должны различаться на  $\pi$ , что осуществляется, когда оптическая разность хода  $\Delta L = 2dn_2$  (см. формулу (29.25)) равна нечетному числу полуволн. Значит, толщина пленки должна удовлетворять условию

$$d = (2m + 1) \frac{\lambda}{4n_2}, \quad m = 0, 1, 2, \dots \quad (32.88)$$

Согласно формуле (32.86) коэффициенты отражения световой волны от передней и задней поверхностей пленки равны соответственно:

$$R_1 = \left( \frac{n_2 - n_1}{n_2 + n_1} \right)^2, \quad R_2 = \left( \frac{n_3 - n_2}{n_3 + n_2} \right)^2. \quad (32.89)$$

Если  $R_1 = R_2$ , то две отраженные волны, находящиеся в противофазе, полностью погасят друг друга и отражения от такого просветленного объектива не возникнет. Условие равенства коэффициентов отражения  $R_1$  и  $R_2$  можно записать в виде

$$(n_2 - n_1)(n_3 + n_2) = (n_3 - n_2)(n_2 + n_1). \quad (32.90)$$

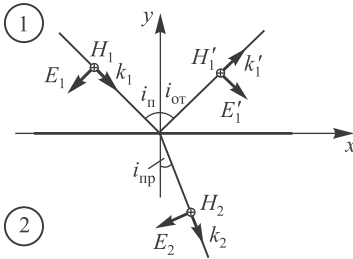
Раскрывая скобки и сокращая одинаковые слагаемые слева и справа, приходим ко второму условию для показателя преломления пленки при однослойном просветлении оптики:

$$n_2 = \sqrt{n_1 n_3}, \quad (32.91)$$



т.е. показатель преломления пленки должен быть равен среднему геометрическому показателей преломления внешней среды и материала линзы. Если внешняя среда — воздух ( $n_1 = 1$ ), а линза изготовлена из флинтowego стекла с показателем преломления  $n_3 = 1,774$ , то условие (32.91) выполняется для покрытия с показателем преломления  $n_2 = 1,33$ . Близкий показатель преломления (1,377) имеет фторид магния  $\text{MgF}_2$ , широко использующийся в оптике (в частности, и для просветления объективов).

### Отражение и преломление $p$ -поляризованной волны



**Рис. 32.7.** Падение  $p$ -поляризованной волны на плоскую границу раздела двух сред: векторы напряженности электрического поля падающей, отраженной и преломленной волн лежат в плоскости падения волны (т.е. в плоскости рисунка), а соответствующие векторы напряженности магнитного поля ортогональны ей, т.е. ортогональны плоскости рисунка и направлены от нас (показано крестиком на луче)

Аналогичным образом рассматривается случай  $p$ -поляризации, когда вектор напряженности электрического поля лежит в плоскости падения волны (рис. 32.7):

$$\begin{aligned} E_{1y} &= -E_1 \sin i_{\text{п}}, & E_{1x} &= -E_1 \cos i_{\text{п}}; \\ E'_{1y} &= -E'_1 \sin i_{\text{от}}, & E'_{1x} &= E'_1 \cos i_{\text{от}}; \\ E_{2y} &= -E_2 \sin i_{\text{пр}}, & E_{2x} &= -E_2 \cos i_{\text{пр}}. \end{aligned} \quad (32.92)$$

Тангенциальные составляющие электрического поля непрерывны на границе раздела двух сред, а нормальные терпят скачок (см. т. 1, разд. 17.6):

$$\begin{aligned} E_{1x} + E'_{1x} &= E_{2x}, \\ \varepsilon_1 (E_{1y} + E'_{1y}) &= \varepsilon_2 E_{2y}. \end{aligned} \quad (32.93)$$

Подставляя сюда выражения (32.92), получаем систему уравнений для амплитуд электрического поля:

$$\begin{aligned} E_1 \cos i_{\text{п}} - E'_1 \cos i_{\text{от}} &= E_2 \cos i_{\text{пр}}, \\ \varepsilon_1 (E_1 \sin i_{\text{п}} + E'_1 \sin i_{\text{от}}) &= \varepsilon_2 E_2 \sin i_{\text{пр}}. \end{aligned} \quad (32.94)$$

Систему уравнений, к которой мы стремимся, проще решать для амплитуд напряженностей магнитного поля. В этих переменных уравнения (32.94) можно переписать как

$$\begin{aligned} \frac{1}{\sqrt{\varepsilon_1}} (H_1 \cos i_{\text{п}} - H'_1 \cos i_{\text{от}}) &= \frac{1}{\sqrt{\varepsilon_2}} H_2 \cos i_{\text{пр}}, \\ \sqrt{\varepsilon_1} (H_1 \sin i_{\text{п}} + H'_1 \sin i_{\text{от}}) &= \sqrt{\varepsilon_2} H_2 \sin i_{\text{пр}}. \end{aligned} \quad (32.95)$$

Третье уравнение вытекает из непрерывности на границе сред тангенциальной составляющей вектора магнитного поля волны, направленного ортогонально координатной плоскости  $xy$  (см. рис. 32.7):

$$H_2 = H_1 + H'_1. \quad (32.96)$$

Подставляя (32.96) в (32.95), приходим к системе уравнений для амплитуд напряженности магнитного поля падающей и отраженной волн:

$$\begin{aligned} H_1 \left( \frac{\cos i_{\text{п}}}{\sqrt{\varepsilon_1}} - \frac{\cos i_{\text{пр}}}{\sqrt{\varepsilon_2}} \right) - H'_1 \left( \frac{\cos i_{\text{от}}}{\sqrt{\varepsilon_1}} + \frac{\cos i_{\text{пр}}}{\sqrt{\varepsilon_2}} \right) &= 0, \\ H_1 (\sqrt{\varepsilon_1} \sin i_{\text{п}} - \sqrt{\varepsilon_2} \sin i_{\text{пр}}) + H'_1 (\sqrt{\varepsilon_1} \sin i_{\text{от}} - \sqrt{\varepsilon_2} \sin i_{\text{пр}}) &= 0. \end{aligned} \quad (32.97)$$

Отличные от нуля решения этой системы существуют только при условии, что ее детерминант равен нулю:

$$\begin{vmatrix} \frac{\cos i_{\text{п}}}{n_1} - \frac{\cos i_{\text{пп}}}{n_2} & -\frac{\cos i_{\text{от}}}{n_1} - \frac{\cos i_{\text{пп}}}{n_2} \\ n_1 \sin i_{\text{п}} - n_2 \sin i_{\text{пп}} & n_1 \sin i_{\text{от}} - n_2 \sin i_{\text{пп}} \end{vmatrix} = 0. \quad (32.98)$$

Как и ранее, мы заменили квадратные корни из диэлектрических проницаемостей на показатели преломления среды. Вычитая второй столбец из первого, приходим к уравнению

$$\begin{vmatrix} \frac{\cos i_{\text{п}} + \cos i_{\text{от}}}{n_1} & -\frac{\cos i_{\text{от}}}{n_1} - \frac{\cos i_{\text{пп}}}{n_2} \\ n_1 (\sin i_{\text{п}} - \sin i_{\text{от}}) & n_1 \sin i_{\text{от}} - n_2 \sin i_{\text{пп}} \end{vmatrix} = 0. \quad (32.99)$$

И в этом случае при  $i_{\text{п}} = i_{\text{от}}$  (угол падения равен углу отражения) и  $n_1 \sin i_{\text{от}} = n_2 \sin i_{\text{пп}}$  (закон Снеллиуса) равна нулю вторая строка детерминанта и вместе с ней весь детерминант. С учетом этих соотношений из (32.97) следует отношение амплитуд отраженной и падающей волн:

$$\frac{H'_1}{H_1} = \frac{E'_1}{E_1} = \frac{n_2 \cos i_{\text{п}} - n_1 \cos i_{\text{пп}}}{n_2 \cos i_{\text{п}} + n_1 \cos i_{\text{пп}}}. \quad (32.100)$$

Теперь легко найти отношение  $H_2/H_1 = 1 + H'_1/H_1$ :

$$\frac{H_2}{H_1} = \frac{2n_2 \cos i_{\text{п}}}{n_2 \cos i_{\text{п}} + n_1 \cos i_{\text{пп}}}. \quad (32.101)$$

Отсюда получаем отношения амплитуд электрических полей:

$$\frac{E_2}{E_1} = \frac{(n_1 H_2)}{(n_2 H_1)} = \frac{2n_1 \cos i_{\text{п}}}{n_2 \cos i_{\text{п}} + n_1 \cos i_{\text{пп}}}. \quad (32.102)$$

Исключая с помощью закона Снеллиуса показатели преломления из (32.100) и (32.102) и применяя затем тригонометрические формулы для синуса двойного угла и суммы (разности) синусов, приходим к выражениям

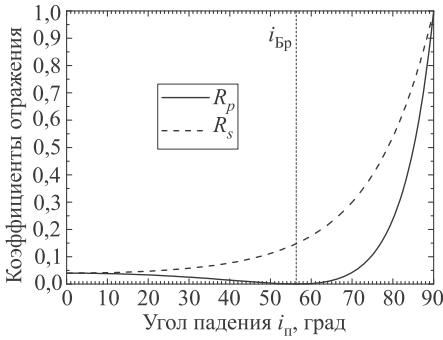
$$\boxed{\begin{aligned} \frac{E'_1}{E_1} &= \frac{\operatorname{tg}(i_{\text{п}} - i_{\text{пп}})}{\operatorname{tg}(i_{\text{п}} + i_{\text{пп}})}, \\ \frac{E_2}{E_1} &= \frac{2 \cos i_{\text{п}} \sin i_{\text{пп}}}{\sin(i_{\text{п}} + i_{\text{пп}}) \cos(i_{\text{п}} - i_{\text{пп}})}. \end{aligned}} \quad (32.103)$$

Коэффициент отражения  $R_p$  для  $p$ -поляризованной волны равен, таким образом,

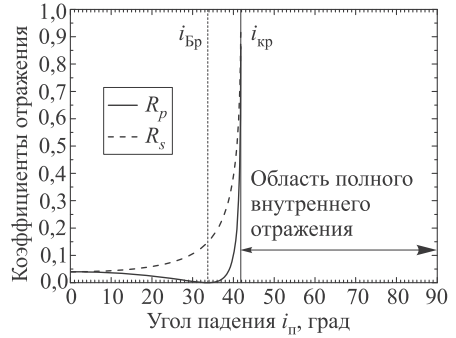
$$\boxed{R_p = \frac{\operatorname{tg}^2(i_{\text{п}} - i_{\text{пп}})}{\operatorname{tg}^2(i_{\text{п}} + i_{\text{пп}})}}. \quad (32.104)$$

При нормальном падении луча из (32.100) и (32.103) следуют для коэффициентов отражения  $R_p$  и прохождения  $D_p = 1 - R_p$  те же самые выражения (32.86) и (32.87), что очевидно, так как в этом случае  $p$ -поляризация не отличается от  $s$ -поляризации.

Формулы (32.84) и (32.103) для отношений амплитуд отраженного и преломленного света к амплитуде падающего света называются **формулами Френеля**. Они были выведены О. Френелем еще в 1821 г. на основе механической модели световых колебаний



**Рис. 32.8.** Коэффициенты отражения  $s$ - и  $p$ -поляризованной волны от стекла с показателем преломления  $n = 1,5$  при падении волны из воздуха; угол Брюстера показан вертикальной пунктирной линией



**Рис. 32.9.** Коэффициенты отражения  $s$ - и  $p$ -поляризованной волны при падении волны из стекла с показателем преломления  $n = 1,5$  в воздух; угол Брюстера показан вертикальной пунктирной линией, критический угол — вертикальной сплошной линией

в упругой среде (эфире). При определенных предположениях они следуют, как мы видели, из классической электродинамики Максвелла.

Если угол падения равен углу Брюстера (см. (31.9)), то отраженный и преломленный лучи взаимно перпендикулярны, т. е.  $i_{\text{пр}} = \pi/2 - i_{\text{п}}$ . В знаменателе коэффициента отражения  $p$ -поляризованной волны (см. (32.104)) оказывается тангенс прямого угла, т. е. бесконечно большое число. Это значит, что  $R_p = 0$ . Иными словами, при падении произвольной волны под углом Брюстера отразится лишь тот компонент, вектор поляризации которого ортогонален плоскости падения. При падении под углом Брюстера коэффициент отражения  $s$ -поляризованной волны, как следует из (32.85), равен:

$$R_s = \cos^2(2i_{\text{п}}) = \left( \frac{1 - \text{tg}^2 i_{\text{п}}}{1 + \text{tg}^2 i_{\text{п}}} \right)^2 = \left( \frac{n^2 - 1}{n^2 + 1} \right)^2. \quad (32.105)$$

Зависимость коэффициентов отражения для волн разных поляризацій показана на рис. 32.8 для  $n_1 = 1$  (воздух) и  $n_2 = 1,5$  (стекло) и для  $n_1 = 1,5, n_2 = 1$  на рис. 32.9. При падении из воздуха в стекло угол Брюстера равен  $i_{\text{Бр}} = \text{arctg} 1,5 = 56,3^\circ$ , а при падении волны из стекла в воздух  $i_{\text{Бр}} = \text{arctg}(1/1,5) = 33,7^\circ$ . Коэффициент отражения при падении под углом Брюстера одинаков в обоих случаях и равен  $R_s = 0,15$ , т. е. отразится лишь 15% полной интенсивности волны. Принципиальная разница между двумя графиками — существование на втором из них области углов, при которых электромагнитная волна вообще не проникает в оптически менее плотную среду. Это происходит при углах падения, больших некоторого критического значения  $i_{\text{кр}}$  (в нашем примере  $i_{\text{кр}} = 41,8^\circ$ ). Это явление называется *полным внутренним отражением*.

## 32.8 Полное внутреннее отражение

С явлением полного внутреннего отражения мы уже знакомы (см. разд. 28.2). При критическом угле падения света из оптически более плотной среды в менее плотную преломленный луч скользит вдоль поверхности раздела. Сейчас нас

интересуют детали того, что происходит, если угол падения превышает критический:  $i_{\text{п}} > i_{\text{кр}}$ . Закон Снеллиуса говорит нам, что в этом случае синус угла преломления должен стать больше единицы, что невозможно, и мы заключаем, что преломленного луча в этом случае не существует. Чтобы глубже разобраться в происходящем, перейдем к рассмотрению явления в терминах волновых векторов.

Модуль волнового вектора равен отношению круговой частоты волны  $\omega$  к скорости ее распространения в среде. Если  $k_0 = \omega/c$  — длина волнового вектора в вакууме, то  $k_1 = k'_1 = k_0 n_1$  — длина волновых векторов в падающей и отраженной волнах, а  $k_2 = k_0 n_2$  — длина волнового вектора в преломленной волне. Проекции этих волновых векторов на поверхность раздела сред (см. рис. 32.6) равны:

$$k_{1x} = k_0 n_1 \sin i_{\text{п}}, \quad k'_{1x} = k_0 n_1 \sin i_{\text{от}}, \quad k_{2x} = k_0 n_2 \sin i_{\text{пр}}. \quad (32.106)$$

Равенство угла отражения углу падения  $i_{\text{п}} = i_{\text{от}}$  и закон Снеллиуса  $n_1 \sin i_{\text{п}} = n_2 \sin i_{\text{пр}}$  эквивалентны равенствам  $k_{1x} = k'_{1x}$  и  $k_{1x} = k_{2x}$  соответственно. Таким образом, тангенциальные составляющие всех трех волновых векторов совпадают.

Нормальную составляющую волнового вектора преломленной волны вычисляем по теореме Пифагора:

$$k_{2y} = -\sqrt{k_2^2 - k_{2x}^2} = -k_0 \sqrt{n_2^2 - n_1^2 \sin^2 i_{\text{п}}}. \quad (32.107)$$

В случае перехода волны из оптически более плотной среды в менее плотную критическому углу  $i_{\text{кр}}$  соответствует обращение в ноль нормальной компоненты волнового вектора. При увеличении угла падения  $i_{\text{п}} > i_{\text{кр}}$  волновой вектор становится мнимой величиной:

$$k_{2y} = -ik_0 \sqrt{n_1^2 \sin^2 i_{\text{п}} - n_2^2}. \quad (32.108)$$

Это означает, что амплитуда преломленной волны убывает по экспоненциальному закону при удалении от поверхности раздела, а глубина проникновения волны в оптически менее плотную среду характеризуется параметром  $\delta$  размерности длины:

$$k_{2y} = -i \frac{1}{\delta}, \quad \delta = \frac{\lambda_0}{\sqrt{n_1^2 \sin^2 i_{\text{п}} - n_2^2}}, \quad (32.109)$$

где  $\lambda_0 = 1/k_0$  — длина падающей волны в вакууме. Действительно, подставляя найденные компоненты волнового вектора в выражение для преломленной волны, получаем

$$\vec{E}_2 = \vec{E}_{02} \exp(i\omega t - ik_2 \cdot \vec{r}) = \vec{E}_{02} e^{-y/\delta} \exp(i\omega t - ik_{2x} x). \quad (32.110)$$

Приведем численный пример: при падении электромагнитной волны с  $\lambda_0 = 580$  нм (желтый цвет) под углом  $i_{\text{п}} = 45^\circ$  из стекла ( $n_1 = 1,5$ ) в воздух ( $n_2 = 1$ ) глубина проникновения  $\delta = 1640$  нм. Иными словами, при полном внутреннем отражении свет все же проникает во вторую среду, но лишь на расстояния порядка его длины волны.

### 32.9 Затухание волн в металле. Скин-эффект

Запишем уравнения Максвелла для однородной изотропной среды с удельной проводимостью  $\sigma$  при отсутствии свободных зарядов. Используя связь  $\vec{j} = \sigma \vec{E}$  плотности тока проводимости с напряженностью электрического поля, имеем

$$\begin{aligned} \operatorname{div} \vec{E} &= 0, \quad \operatorname{div} \vec{B} = 0, \\ \operatorname{rot} \vec{E} &= -\frac{\partial \vec{B}}{\partial t}, \quad \operatorname{rot} \vec{B} = \mu\mu_0\sigma \vec{E} + \frac{\mu\varepsilon}{c^2} \frac{\partial \vec{E}}{\partial t}. \end{aligned} \quad (32.111)$$

Волновое уравнение для полей мы найдем, применяя к левым частям уравнений в нижней строчке (32.111) операцию ротора и используя известную формулу векторного анализа  $\operatorname{rot} \operatorname{rot} \vec{a} = \operatorname{grad} \operatorname{div} \vec{a} - \Delta \vec{a}$ . Поскольку дивергенции наших полей равны нулю, для электрического поля мы получаем волновое уравнение

$$\Delta \vec{E} - \frac{\mu\varepsilon}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} - \mu\mu_0\sigma \frac{\partial \vec{E}}{\partial t} = 0. \quad (32.112)$$

Точно так же выглядит уравнение для магнитной индукции  $\vec{B}$ .

Решение по-прежнему ищем в виде плоской волны

$$\vec{E}(t, \vec{r}) = \vec{E}_0 \exp(i\omega t - i\vec{k} \cdot \vec{r}). \quad (32.113)$$

Подставляя это выражение в (32.112), получаем уравнение для длины волнового вектора  $\vec{k}$

$$k^2 = \frac{\mu\varepsilon\omega^2}{c^2} - i\mu\mu_0\sigma\omega = 0. \quad (32.114)$$

Очевидно, что  $k$  будет комплексным числом, но мы уже знаем, что это означает экспоненциальное затухание волны. Вводим показатель преломления  $n$  проводящей среды и коэффициент затухания  $\gamma$ :

$$k = \frac{\omega}{c}(n - i\gamma), \quad (32.115)$$

где  $\omega/c$  — длина волнового вектора в вакууме. Подставляя (32.115) в (32.114) и приравнявая по отдельности действительные и мнимые части слева и справа от знака равенства, получаем уравнения

$$\begin{aligned} n^2 - \gamma^2 &= \mu\varepsilon, \\ n\gamma &= \frac{\mu\mu_0\sigma c^2}{2\omega} = \frac{\mu\sigma}{2\varepsilon_0\omega}. \end{aligned} \quad (32.116)$$

Умножим первое уравнение на  $n^2$ :

$$n^4 - \mu\varepsilon n^2 - (\gamma n)^2 = 0. \quad (32.117)$$

Подставим в третий член выражение для  $\gamma n$  из (32.116):

$$n^4 - \mu\varepsilon n^2 - \left(\frac{\mu\sigma}{2\varepsilon_0\omega}\right)^2 = 0. \quad (32.118)$$

Решение этого биквадратного уравнения дает нам показатель преломления проводящей среды:

$$n = \sqrt{\frac{\varepsilon\mu}{2}} \sqrt{1 + \sqrt{1 + \left(\frac{\sigma}{\varepsilon_0\varepsilon\omega}\right)^2}}. \quad (32.119)$$

Отсюда следует выражение для коэффициента поглощения

$$\gamma = \frac{\mu\sigma}{2\varepsilon_0\omega n} = \sqrt{\frac{\varepsilon\mu}{2}} \sqrt{-1 + \sqrt{1 + \left(\frac{\sigma}{\varepsilon_0\varepsilon\omega}\right)^2}}. \quad (32.120)$$

При низкой проводимости среды ( $\sigma \ll \varepsilon_0\varepsilon\omega$ ) показатель преломления имеет стандартный вид  $n \approx \sqrt{\varepsilon\mu}$ , а коэффициент поглощения

$$\gamma \approx \frac{\sigma}{2\varepsilon_0\varepsilon\omega} \sqrt{\varepsilon\mu}. \quad (32.121)$$

При высокой проводимости среды ( $\sigma \gg \varepsilon_0\varepsilon\omega$ )

$$n \approx \gamma \approx \sqrt{\frac{\mu\sigma}{2\varepsilon_0\omega}}. \quad (32.122)$$

Если волна с комплексным волновым вектором распространяется в положительном направлении оси  $x$ , то ее амплитуда убывает по закону  $E_0 \exp(-\omega\gamma x/c)$ , т. е. падает в  $e = 2,7$  раза на расстоянии  $\delta = c/(\omega\gamma)$ . Поскольку интенсивность волны пропорциональна квадрату электрического поля, закон убывания интенсивности выглядит как  $I(x) = I_0 \exp(-2x/\delta)$ . Параметр  $\delta$  называют **глубиной проникновения**. При высокой проводимости глубина проникновения равна:

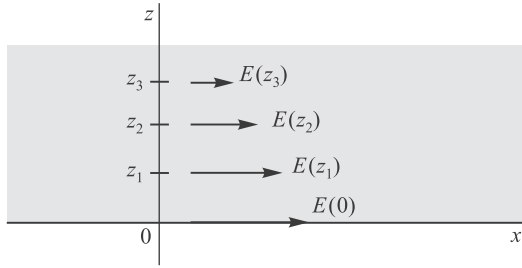
$$\delta \approx \sqrt{\frac{2}{\mu\mu_0\sigma\omega}}. \quad (32.123)$$

Это означает, что при падении электромагнитной волны на среду с высокой проводимостью коэффициент отражения близок к единице. Поэтому, например, металлы блестят — они хорошо отражают видимый свет.

Чем выше частота волны, тем на меньшее расстояние проникает она в проводящую среду. То же самое можно сказать о переменном токе в проводнике, толщиной которого нельзя пренебречь, как мы это делали раньше. Сейчас нас интересует распределение тока в толще проводника и мы рассмотрим идеализированную систему, для которой решение выглядит наиболее просто. Пусть проводник занимает полупространство  $z > 0$ , а ток течет в положительном направлении оси  $Ox$ .

Уравнения Максвелла в однородной проводящей среде имеют вид

$$\begin{aligned} \operatorname{div} \vec{E} &= 0, & \operatorname{div} \vec{B} &= 0, \\ \operatorname{rot} \vec{E} &= -\frac{\partial \vec{B}}{\partial t}, & \operatorname{rot} \vec{B} &= \mu\mu_0\sigma \vec{E}. \end{aligned} \quad (32.124)$$



**Рис. 32.10.** Полупространство  $z > 0$  заполнено проводником, в котором под действием электрического поля  $E(z)$  волны в направлении оси  $Ox$  течет переменный ток; как поле, так и плотность тока экспоненциально затухают в толще проводника по мере удаления от его поверхности

Здесь мы пренебрегли током смещения по сравнению с током проводимости. Применяя, как и ранее, операцию ротор к выражениям во второй строчке (32.124), приходим к волновому уравнению для электрического поля:

$$\Delta \vec{E} = \mu\mu_0\sigma \frac{\partial \vec{E}}{\partial t}. \quad (32.125)$$

По условию ток направлен вдоль оси  $Ox$ . По закону Ома так же должно быть направлено электрическое поле волны. Вследствие симметрии задачи амплитуда волны может зависеть лишь от координаты  $z$ , т.е. от расстояния до края проводника (рис. 32.10):

$$\vec{E} = \vec{n}_x E(z) e^{i\omega t}. \quad (32.126)$$

Подставляем это выражение в уравнение для электрического поля (32.125):

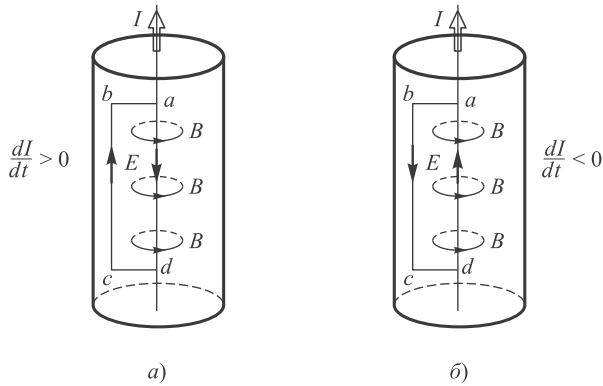
$$\frac{d^2 E(z)}{dz^2} = i\mu\mu_0\sigma\omega E(z). \quad (32.127)$$

Убывающее при больших  $z$  решение имеет вид

$$E(z) = E(0)e^{-z/\delta} \exp i(\omega t - z/\delta). \quad (32.128)$$

По такому же закону изменяется и плотность электрического тока  $\vec{j} = \sigma \vec{E}$ . Мы видим, что глубина проникновения переменного тока в толщу проводника совпадает с глубиной проникновения электромагнитной волны в среду с высокой проводимостью — (см. (32.123) и рис. 32.10). В пределе бесконечно большой проводимости (идеальный проводник) толщина слоя стремится к нулю — поле не проникает в проводник. Это явление называется *скин-эффектом* (от англ. *skin* — кожа). При увеличении частоты поля скин-слой также уменьшается. Наоборот, при стремлении частоты к нулю скин-слой увеличивается до бесконечности: постоянный ток распределен по проводнику равномерно.

Джоулево тепло выделяется в приповерхностном слое. Толщина скин-слоя меняется в широких пределах. В меди при частоте тока 50 Гц его толщина равна 9,34 мм, для тока высокой частоты 100 кГц она составляет уже 0,21 мм.



**Рис. 32.11.** При возрастании тока через проводник ( $dI/dt > 0$ ) индукционный ток в контуре  $abcd$  течет по часовой стрелке, при убывании ( $dI/dt < 0$ ) — против; и в том, и в другом случае ЭДС индукции препятствует изменению тока на оси проводника и способствует его изменению на периферии, что приводит к скин-эффекту

Скин-эффект позволяет применять полые кабели или кабели, покрытые слоем металла с особенно высокой проводимостью, что снижает расход материалов и мощностей. Вредные последствия скин-эффекта — возрастание сопротивления за счет уменьшения эффективной площади поперечного сечения проводника.

Качественно картину возникновения скин-эффекта можно понять с помощью правила Ленца. Если по проводу течет ток  $I$ , то вокруг него имеется круговое магнитное поле  $B$  (рис. 32.11). Выделим в толще проводника контур  $abcd$ . Магнитный поток через него направлен ортогонально плоскости рисунка на нас. При увеличении тока ( $dI/dt > 0$ ) этот магнитный поток возрастает, что приводит к появлению кругового электрического поля  $E$ . Порождаемый им индукционный ток создает магнитное поле, препятствующее росту магнитного потока, т. е. направленное от нас. Значит, электрическое поле индукции направлено в этом случае по часовой стрелке (рис. 32.11, *a*). В обратном случае уменьшения тока ( $dI/dt < 0$ ) электрическое поле индукции направлено против часовой стрелки (рис. 32.11, *б*). И в том, и в другом случае электромагнитная индукция препятствует изменению тока на оси проводника и способствует изменению тока на периферии. Чем дальше от оси провода, тем больше амплитуда колебаний тока, так как растет площадь, охватываемая контуром  $abcd$ . В результате средняя плотность тока вблизи поверхности проводника выше, чем на его оси. Чем быстрее меняется ток, тем сильнее выражен скин-эффект.

## 32.10 Дипольное излучение

Излучение электромагнитных волн движущимися зарядами, как и все другие явления классической электродинамики, описывается уравнениями Максвелла. Однако их решения, требующие свободного владения достаточно сложным математическим аппаратом, обычно рассматриваются уже в рамках курса теоретической физики. Все же нельзя совсем обойти своим вниманием столь важный



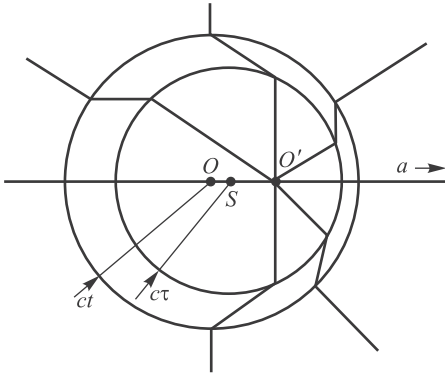
раздел физики, имеющий огромное практическое значение. Современный мир буквально погружен в электромагнитные волны, излучаемые многочисленными антеннами теле- и радиостанций, ретрансляторами сотовой связи, да и нашими бытовыми приборами. Поэтому в данном разделе мы попытаемся изложить основные черты теории излучения «на пальцах», не прибегая к сложным формулам.

Простейшая излучающая система — электрический диполь. Уравнения Максвелла приводят к заключению, что электромагнитные волны излучаются системой зарядов, если ее дипольный момент меняется во времени (более точно, если отлична от нуля *вторая производная* дипольного момента). Мы подойдем к этому результату, рассматривая ускоренное движение электрического заряда. Понятно, что покоящийся точечный заряд излучать не должен — откуда бы взяться энергии, уносимой излучением? Столь же понятно, что равномерное прямолинейное движение заряда в вакууме также не должно порождать излучение в силу принципа относительности, потому что всегда можно перейти в систему отсчета, где этот заряд покоится. Поэтому к проблеме излучения мы подойдем, рассматривая *ускоренное* движение точечного заряда. Чтобы ускорить заряд, мы должны совершить работу, и часть этой работы пойдет на излучение энергии в форме электрического и магнитного полей. Напряженности этих полей несут информацию о внезапном изменении скорости заряда.

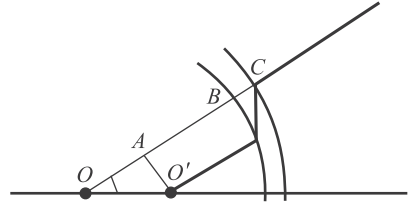
Профессор Университета штата Нью-Йорк в Стоуни-Брук Кл. Э. Суорц<sup>2</sup> рекомендовал в качестве подготовки вспомнить движение струи, вырывающейся из расположенного горизонтально узкого наконечника шланга. Поскольку нас не интересует падение струи по параболе вследствие земного тяготения, мы должны рассматривать лишь ее горизонтальную проекцию. При неподвижном наконечнике прямая линия струи моделирует силовую линию электрического поля, выходящего из заряда. Если наконечник медленно и равномерно двигать в направлении, перпендикулярном струе, то струя будет следовать за ним, сохраняя прямолинейную форму. Если наконечник двигать с ускорением, струя начнет от него отставать и на ней образуется излом в направлении, перпендикулярном основному движению воды в струе. Если, наконец, двигать наконечником вправо — влево, то по струе будет перемещаться извивающаяся волна, размах которой увеличивается с увеличением расстояния до наконечника.

Вооружившись знанием этого явления, хорошо известного нам из повседневного опыта, мы переходим к рассмотрению электрического поля ускоряющегося заряда, следуя книге Суорца. Процесс схематически изображен на рис. 32.12. Пусть точечный заряд  $q$  покоится в точке  $O$ , которую мы выбираем за начало координат. В момент времени  $t = 0$  на заряд начинает действовать направленная вправо внешняя сила, сообщающая ему ускорение  $a$ . Через время  $\Delta t$  сила прекращает свое действие, заряд к этому моменту находится в точке  $S$  и продолжает движение вправо с постоянной скоростью  $v = a\Delta t$ . Рассмотрим картину силовых линий электрического поля через время  $\tau$  после начала равномерного движения заряда (т. е. к моменту наблюдения  $t = \tau + \Delta t$ ), когда заряд займет положение  $O'$ . Электромагнитное поле, несущее информацию об изменении зарядом положения

<sup>2</sup>Кл. Э. Суорц. Необыкновенная физика обыкновенных явлений. Т. 2. М.: Наука, 1987.



**Рис. 32.12.** Силовые линии электрического поля заряда  $q$ , ускорившегося от точки  $O$  до точки  $S$  в течение времени  $\Delta t$ ; в момент наблюдения  $t = \tau + \Delta t$  заряд находится в точке  $O'$



**Рис. 32.13.** К вычислению поперечной составляющей электрического поля излучения: компонента электрического поля, ортогональная направлению распространения волны, возникает вследствие излома силовых линий в слое между сферой радиусом  $ct$  с центром в  $O$  и сферой радиусом  $ct$  с центром в точке  $S$ , которую мы считаем близкой к точке  $O$

в пространстве, распространяется в вакууме со скоростью света  $c$ . Проведем сферу радиусом  $ct$  с центром в точке  $O$  первоначального местонахождения заряда. За ее пределами электрические силовые линии направлены по радиусам от точки  $O$ , поскольку информация о начале движения заряда за время  $t$  туда достичь еще не успела. Теперь проведем сферу радиусом  $ct$  с центром в точке  $S$  местонахождения заряда к моменту прекращения действия на него ускоряющей силы. Поле внутри этой сферы — это поле равномерно движущегося заряда. Если время  $\Delta t \ll t$  и заряд движется медленно, то к моменту наблюдения  $t$  силовые линии электрического поля направлены по радиусам, исходящим из точки  $O'$ .

Таким образом, поле излучения, несущее информацию об ускоренном движении заряда, сосредоточено в слое между двумя сферами. Сфера радиусом  $ct$  с центром в точке  $O$  представляет собой передний край электромагнитного импульса в момент времени  $t$ . Сфера радиусом  $ct$  с центром в точке  $S$  — это задний край электромагнитного импульса. Со временем сферы расширяются и объем слоя увеличивается — излучение распространяется в пространстве. Как видно из рисунка 32.12, силовые линии имеют в этом слое излом, так что помимо обычной продольной компоненты (статическое поле, направленное по радиусам) появляется и поперечная компонента электрического поля, ортогональная радиусам сфер, т. е. направлению распространения излучения.

Найдем эту компоненту поля (рис. 32.13). При  $\Delta t \ll t$  расстояние  $OO' = v\tau = a\tau\Delta t$ . Тогда расстояние между лучами равно  $O'A = OO' \sin \theta \approx a\tau\Delta t \sin \theta$ . Расстояние между сферами  $BC \approx c\Delta t$ . Отношение  $O'A/BC$  равно отношению

продольной и поперечной составляющих электрического поля:

$$\frac{E_{\perp}}{E_{\parallel}} = \frac{a\tau \sin \theta}{c}. \quad (32.129)$$

Вводя обычное обозначение для расстояния  $r = ct \approx c\tau$  от заряда до точки наблюдения, записываем (32.129) в виде

$$E_{\perp} = E_{\parallel} \frac{ar \sin \theta}{c^2}. \quad (32.130)$$

Продольная составляющая электрического поля выражается законом Кулона:

$$E_{\parallel} = \frac{q}{4\pi\epsilon_0 r^2}. \quad (32.131)$$

Подстановка (32.131) в (32.130) дает нам искомое выражение для поперечной составляющей электрического поля излучения:

$$E_{\perp} = \frac{aq \sin \theta}{4\pi\epsilon_0 c^2 r}. \quad (32.132)$$

Как и ожидалось, поперечная составляющая электрического поля зависит от ускорения заряда. Но самое замечательное состоит в том, что она обратно пропорциональна *первой степени расстояния* в отличие от статического поля, убывающего как  $1/r^2$ . На больших расстояниях от заряда статическое поле становится пренебрежимо малым по сравнению с полем излучения. Поток энергии излучения выражается квадратом амплитуды поля. Соответственно он убывает по закону обратных квадратов, так что его произведение на площадь сферы вокруг заряда остается постоянной, как и следует ожидать из закона сохранения энергии. Именно такую зависимость от расстояния (типа (32.132)) имеет поле (27.71) сферической волны. Следует отметить, что всплеск электрического поля, порожденный кратковременным ускорением электрического заряда, сопровождается связанным с ним всплеском магнитного поля, ортогонального как электрическому полю, так и направлению распространения излучения.

Еще одним замечательным свойством выведенного соотношения (32.132) является зависимость электрического поля излучения от угла  $\theta$ : максимальная интенсивность излучения приходится на направления, ортогональные ускорению ( $\theta = 90^\circ$ ); в направлении же вдоль ускорения ( $\theta = 0^\circ, 180^\circ$ ) заряд не излучает. Возникает своего рода парадокс: излучаемая сферическая электромагнитная волна (ее волновой фронт представляет собой сферы) не обладает сферической симметрией, поскольку ее интенсивность различна в разных направлениях.

В окрестности любой точки наблюдения на больших расстояниях от ускоряемого заряда (в так называемой *волновой зоне*) сфера мало отличается от плоскости, и мы можем применить соотношение (32.38), связывающее амплитуды колебаний электрического и магнитного полей в плоской волне. Отсюда мы находим амплитуду магнитного поля излучения ускоряемого заряда

$$H_{\perp} = \sqrt{\frac{\epsilon_0}{\mu_0}} E_{\perp} = \epsilon_0 c E_{\perp} \quad (32.133)$$

и по формуле (32.65) получаем выражение для плотности потока энергии излучения

$$S = E_{\perp} H_{\perp} = \frac{1}{4\pi\epsilon_0 c^3} \frac{(qa)^2}{4\pi_0 r^2} \sin^2 \theta. \quad (32.134)$$

Умножая  $S$  (абсолютную величину вектора Пойнтинга) на элемент площади сферической поверхности  $d\Sigma = r^2 \sin\theta d\theta d\varphi$  и интегрируя по всей сфере радиусом  $r$ , получаем **мощность излучения** — количество энергии, излучаемой в единицу времени:

$$N = \oint S d\Sigma = \frac{1}{4\pi\epsilon_0 c^3} \frac{(qa)^2}{4\pi_0 r^2} r^2 \int_0^{\pi} \sin^3 \theta d\theta \int_0^{2\pi} d\varphi = \frac{(qa)^2}{8\pi\epsilon_0 c^3} \int_0^{\pi} \sin^3 \theta d\theta. \quad (32.135)$$

Делая в интеграле замену переменных  $x = \cos\theta$ , легко вычисляем его:

$$\int_0^{\pi} \sin^3 \theta d\theta = \int_{-1}^1 (1-x^2) dx = \frac{4}{3}. \quad (32.136)$$

Подставляя (32.136) в (32.135), приходим к **формуле Лармора** для мощности излучения заряда  $q$ , движущегося с ускорением  $a$ :

$$N = \frac{2}{3} \frac{(qa)^2}{4\pi\epsilon_0 c^3}. \quad (32.137)$$

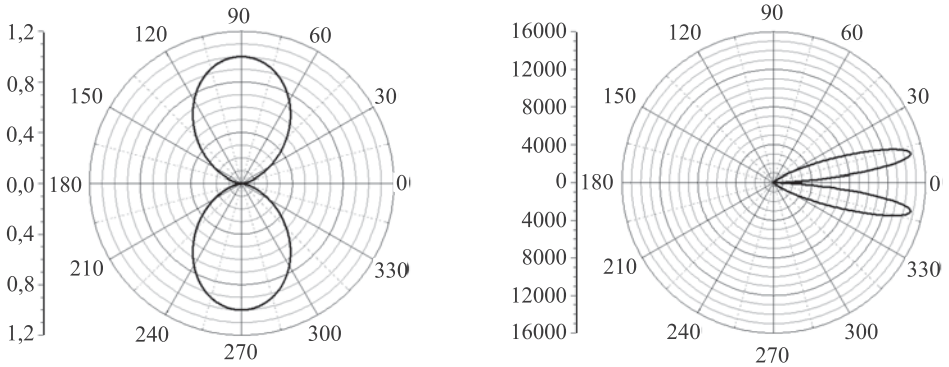
Мы пришли к выводу, что заряженные частицы, движущиеся в вакууме, излучают только при ускоренном движении. Так, при изменении направления движения заряда в электрическом поле (например, при рассеянии электрона на протоне) возникает **тормозное излучение**. Ускорение заряженной частицы обратно пропорционально ее массе, поэтому мощность тормозного излучения обратно пропорциональна квадрату массы частицы. Отсюда следует, что тормозное излучение при рассеянии электронов в миллионы раз сильнее аналогичного излучения для протонов. Тормозное излучение, возникающее при резком торможении (соударении) электронов с анодом, является одним из источников рентгеновского излучения в катодных трубках. Тормозное излучение возникает также при столкновении заряженных частиц в горячей плазме. Подобный тепловой механизм частично ответствен за космическое рентгеновское излучение, наблюдаемое с помощью рентгеновских телескопов, выводимых на орбиту вокруг Земли.

Применим формулу Лармора для заряда, колеблющегося по гармоническому закону с амплитудой  $x_0$  и круговой частотой  $\omega$ :

$$x(t) = x_0 \cos \omega t, \quad v(t) = \dot{x}(t) = -x_0 \omega \sin \omega t, \quad a = \ddot{x}(t) = -x_0 \omega^2 \cos \omega t. \quad (32.138)$$

Тогда мощность излучения в момент времени  $t$  выражается как

$$N(t) = N = \frac{2}{3} \frac{(qx_0)^2 \omega^4}{4\pi\epsilon_0 c^3} \cos^2 \omega t. \quad (32.139)$$



**Рис. 32.14.** Отношение  $S(\theta)/S(\pi/2)$  потока энергии при произвольном угле  $\theta$  между направлениями излучения и ускорения заряда к потоку энергии, излучаемой в поперечном ( $\theta = \pi/2$ ) направлении, при разных скоростях заряда:  $v = 0$  для левой диаграммы и  $v = 0,9c$  для правой

Средняя мощность излучения за период  $T = 2\pi/\omega$  определяется интегрированием (32.139) по времени:

$$\langle N \rangle = \frac{1}{T} \int_0^T N(t) dt = \frac{1}{3} \frac{(qx_0)^2 \omega^4}{4\pi\epsilon_0 c^3}. \quad (32.140)$$

Зависимость мощности излучения от четвертой степени частоты — это уже упоминавшийся в разделе 31.4 закон Рэлея, объясняющий голубой цвет неба и красный цвет солнца на закате.

Электрический дипольный момент заряда равен  $\vec{p} = q\vec{r}$ , поэтому произведение заряда на ускорение в формуле Лармора можно представить как вторую производную дипольного момента:  $(qa)^2 = \ddot{\vec{p}}^2$ . В таком виде выведенные соотношения пригодны для описания излучения любой системы зарядов. Кроме того, в аргументе дипольного момента следует учесть запаздывание: поле излучения, наблюдаемое в момент времени  $t$ , создано движением заряда в более ранний момент времени  $t - r/c$ , где  $r$  — расстояние от заряда до точки наблюдения. Тогда формулу (32.134) можно записать в виде

$$S = \frac{1}{4\pi\epsilon_0 c^3} \frac{|\ddot{\vec{p}}(t - r/c)|^2}{4\pi_0 r^2} \sin^2 \theta. \quad (32.141)$$

Вывод о преимущественном излучении ускоренного заряда в поперечном направлении справедлив лишь для медленных зарядов. Так излучают, например, линейные антенны, переменный ток в которых и означает колебание зарядов. Но если ускоряемый заряд движется с релятивистской скоростью, то **диаграмма направленности** дипольного излучения (зависимость интенсивности от направления излучения) меняет свой вид, и максимальная энергия излучается в направлении скорости. Это иллюстрируется на диаграммах рис. 32.14, где показана

относительная величина потока энергии для частного случая равноускоренного движения, когда ускорение и скорость заряда параллельны. На обеих диаграммах ускорение заряда направлено вправо вдоль линии  $180 - 0$  градусов. Разница в том, что диаграмма слева соответствует моменту, когда скорость заряда  $v = 0$  (или очень мала). В этом случае максимум излучения направлен ортогонально ускорению, что мы уже нашли ранее. На правом же рисунке показана диаграмма направленности излучения заряда, скорость которого достигла величины  $v = 0,9c$ . В этом случае максимальная энергия излучается в направлении, близком к направлению скорости заряда, независимо от направления ускорения.

Видно, что помимо изменения диаграммы направленности с ростом скорости из-за эффекта Доплера резко возрастает и поток энергии в максимуме излучения.

Для угла  $\theta_{\max}$ , на который приходится максимальное излучение, уравнения Максвелла дают выражение

$$\theta_{\max} = \arccos \left( \frac{\sqrt{1 + 15\beta^2} - 1}{3\beta} \right), \quad \beta = v/c. \quad (32.142)$$

Для медленного заряда ( $v \sim 0$ ) имеем максимум излучения в направлении  $\theta_{\max} = 90^\circ$ , ортогональном ускорению. При  $v = 0,5c$  получаем  $\theta_{\max} = 38,2^\circ$ , а для скорости  $v = 0,9c$  отсюда следует значение  $\theta_{\max} = 13,4^\circ$  (см. рис. 32.14). Учитывая соотношение  $E = mc^2/\sqrt{1 - \beta^2}$  между энергией и скоростью частицы (см. формулу (9.67) в т. 1), легко получить из (32.142), что для ультрарелятивистских частиц ( $E \gg mc^2$ ) излучение сосредоточено в узком конусе с углом раствора  $2\theta_{\max} \approx mc^2/E$  — так называемый *прожекторный эффект*.

Этот эффект характерен для *синхротронного (магнито-тормозного) излучения* в рентгеновском диапазоне, испускаемого электронами, движущимися по круговым траекториям в магнитном поле ускорителей. При малых энергиях излучение вращающегося электрона происходит на циклотронной частоте, определяемой периодом вращения (см. формулу (19.9) в т. 1). При увеличении энергии электрона излучение смещается в сторону более высоких частот и распределяется среди большего числа высших гармоник, спектральные линии которых расположены очень близко друг к другу. Так как каждая из них имеет конечную ширину, спектральные линии перекрываются, так что спектр синхротронного излучения является непрерывным. Сильнейшим источником синхротронного излучения является пульсар (нейтронная звезда) в центре Крабовидной туманности, образовавшейся на месте взрыва сверхновой в 1054 г. Вспышка была видна три недели даже в дневное время и зафиксирована арабскими и китайскими астрономами. Синхротронное излучение применяется в медицине и материаловедении при изучении физико-химических свойств твердых тел.

## 32.11 Излучение равномерно движущегося заряда

Излучение ускоренно движущегося заряда часто описывают как «сбрасывание» электромагнитного поля при резкой остановке (или резком старте) заряженной частицы. Если более основательно задуматься над причиной возникновения

этого излучения, станет ясно, что всему виной уже знакомый нам принцип Гюйгенса. Каждая точка пространства, мимо которой движется заряд, может рассматриваться как источник вторичных волн, которые интерферируют между собой. При равномерном движении эти вторичные волны взаимно гасят друг друга и излучение не возникает. При ускоренном движении электрона силовые линии поля в одних местах сгущаются, в других разрежаются, возникает уже обсуждавшийся излом силовых линий, т. е. дипольное излучение.

Однако вывод о взаимном гашении вторичных волн относится лишь к движению заряда в вакууме. Фазовая скорость света в среде с показателем преломления  $n$  равна  $c/n$ . Если электрон движется с постоянной скоростью  $v_e > c/n$ , то возникает ударная световая волна, подобная акустической ударной волне при движении объекта со сверхзвуковой скоростью. Как мы видели в разделе 27.5, фронт ударной волны формируется в таком случае как огибающая вторичных волн, т. е. вследствие их интерференции согласно принципу Гюйгенса. Электрон, обгоняющий свет, излучает энергию под углом  $\varphi$  к направлению движения, причем  $\cos \varphi = c/(nv_e)$ . Это излучение называется **излучением Черенкова**. Точно так же как для звука образовывался конус Маха, при движении быстрого электрона образуется **черенковский конус** с углом раствора  $\theta$  таким, что  $\sin \theta = c/(nv_e)$ . Эти формулы аналогичны полученным ранее в разделе 27.5, только скорость звука  $v$  заменена здесь на скорость света  $c/n$  в данной среде. Излучение «сверхсветового» электрона было обнаружено П. А. Черенковым (1934), теоретическое объяснение дано И. Е. Таммом и И. М. Франком (1937). Они стали первыми отечественными учеными, удостоенными Нобелевской премии по физике<sup>3</sup>, присужденной им в 1958 г. «за открытие и объяснение эффекта Черенкова»<sup>4</sup>. На основе этого эффекта работают широко используемые в физике черенковские счетчики, позволяющие по углу раствора конуса Маха определить скорость частицы.

Излучение Черенкова относится к равномерному движению «сверхсветового» электрона в бесконечной однородной среде. Но бесконечных сред не бывает, электроны могут влетать в среду и вылетать из нее. Что происходит при пересечении электроном границы раздела двух сред? Этим вопросом задались в 1944 г. советские физики В. Л. Гинзбург и И. М. Франк и открыли на кончике пера **переходное излучение**, возникающее при пересечении равномерно движущимися заряженными частицами границы раздела между средами с разными показателями преломления. Первые (и весьма несовершенные) экспериментальные результаты были получены лишь в 1959 г.

Суть эффекта в том, что фазовые скорости света различаются в разных средах и при пересечении зарядом границы между ними вторичные волны не

<sup>3</sup>За два года до этого, в 1956 г., Нобелевская премия по химии была присуждена Н. Н. Семенову «за исследования в области механизма химических реакций».

<sup>4</sup>В отечественной литературе это явление называется эффектом Вавилова—Черенкова. П. А. Черенков работал в лаборатории крупнейшего советского ученого, президента АН СССР С. И. Вавилова, который инициировал эти исследования. Справедливости ради заметим, что еще в 1904 г. А. Зоммерфельд рассчитал подобное излучение для электрона, движущегося в вакууме со сверхсветовой скоростью. Созданная вскоре специальная теория относительности сделала его расчеты бессмысленными, как тогда казалось, и работа была забыта. О применении расчетов Зоммерфельда к движению электрона в среде никто в то время не подумал.

гасят друг друга даже при досветовых скоростях  $v_e < c/n$  заряда. Так будет происходить всегда при движении заряда в среде с показателем преломления, меняющимся в пространстве (или во времени). Причина, в сущности, та же, что и при возникновении дипольного излучения: из-за разной скорости света в разных точках траектории заряда вторичные волны, испущенные в один момент времени, будут отставать (опережать) вторичные волны, испущенные в другой момент. В результате их интерференции образуется фронт волны излучения. Гинзбург и Франк рассмотрели падение электрона на плоскую металлическую поверхность. Поле заряда электрона и наведенных положительных зарядов, распределенных по металлической поверхности, совпадает с полем двух точечных зарядов: электрона и его «положительного отражения» в металлическом зеркале. Когда электрон приближается к металлу, его «изображение» движется навстречу ему, дипольный момент этой пары меняется, но равномерно, и поэтому излучения не возникает. Но при пересечении границы металл — вакуум электрон попадает в окружение многих электронов проводимости и «теряется» среди них. Выглядит это так, словно заряды электрона и его «изображения» аннигилируют, исчезают. Это почти мгновенное исчезновение дипольного момента равносильно резкой остановке двух зарядов, что и приводит к всплеску переходного излучения. Такое простое и наглядное объяснение на пальцах было дано Гинзбургом в научно-популярной статье<sup>5</sup>. Оно позволяет понять, почему переходное излучение может возникнуть не только при пересечении зарядом границы раздела сред, но и при пролете электрона вблизи металлических поверхностей с меняющимся профилем (шары, гофрированные поверхности, всякого рода металлические углы и т. п.). В этих случаях скорость изменения дипольного момента, образованного электроном и его «изображением» в находящейся поблизости среде, уже не постоянна, что и является причиной возникновения переходного излучения.

## Контрольные вопросы

1. Как направлены поля в плоской электромагнитной волне?
2. Чем различаются эффекты Доплера для электромагнитной и звуковой волн?
3. От чего зависит интенсивность электромагнитной волны?
4. Почему свет оказывает давление?
5. Что такое полное внутреннее отражение и критический угол падения?
6. Что происходит, если угол падения света превышает критическое значение?
7. Во сколько раз уменьшится интенсивность электромагнитной волны на расстоянии, равном двум глубинам проникновения?
8. Что такое скин-эффект?
9. Что такое тормозное излучение?
10. Какова диаграмма направленности дипольного излучения медленного заряда?

---

<sup>5</sup> В. Л. Гинзбург. Переходное излучение и переходное рассеяние. Природа, № 8, 1975, с. 56.



11. Как меняется диаграмма направленности дипольного излучения при ускорении заряда, движущегося с релятивистской скоростью?
12. Опишите возникновение и особенности синхротронного излучения.
13. В каком случае происходит излучение равномерно движущегося заряда? Каковы особенности этого излучения?
14. Что такое переходное излучение и каковы причины его возникновения?

## Часть VI

# ОСНОВЫ КВАНТОВОЙ МЕХАНИКИ

*Быть может, эти электроны —  
Миры, где пять материков,  
Искусства, знания, войны, троны  
И память сорока веков.  
Еще, быть может, каждый атом —  
Вселенная, где сто планет;  
Там все, что здесь, в объеме сжатом,  
Но также то, чего здесь нет. . .*

---

В. Брюсов. «Мир электрона»



## Глава 33

# Квантовая природа излучения

Два века борьбы корпускулярной и волновой теорий света завершились победой последней. Уравнения Максвелла стали венцом теоретической физики XIX в. На горизонте науки виднелись лишь два темных пятнышка — проблемы с тепловым излучением и результат опыта Майкельсона. Казалось, что еще немного, и физики останутся без работы: законы природы представлялись в основном понятными. К счастью, так не случилось. Из одного пятнышка выросла квантовая теория, из другого — теория относительности. С теорией относительности мы познакомились в первом томе курса. Настало время приоткрыть дверь в квантовый мир. Первым это сделал немецкий ученый Макс Планк, удостоенный в 1918 г. Нобелевской премии «в знак признания заслуг в развитии физики вследствие открытия им квантов энергии». Макс Планк сделал свой доклад о квантовой теории теплового излучения на заседании Немецкого физического общества 14 декабря 1900 г. Вступая в XX век, мир, того не подозревая, вступил в квантовую эпоху.

### 33.1 Тепловое излучение

*Тепловое излучение* — это электромагнитное излучение широкого спектрального состава, испускаемое телами и обусловленное тепловыми колебаниями электрических зарядов (электронов и ионов), входящих в состав вещества.

Поэтому характеристики теплового излучения (его интенсивность и спектральный состав) зависят от температуры тела. Все прочие виды электромагнитного излучения существуют за счет других, нетепловых форм энергии. Тепловое излучение — единственный вид излучения, которое может находиться в термодинамическом равновесии с излучающими телами.

Предположим, что нагретое тело помещено в полость с идеально отражающими стенками, которые поддерживаются при некоторой постоянной температуре  $T$ . Если в полости нет никакой среды (газа), то обмен энергией между оболочкой и телом происходит только за счет теплового излучения. С течением времени температура тела станет равной температуре оболочки и наступит динамическое равновесие — в единицу времени тело будет поглощать столько же энергии,

сколько и излучать. Допустим, что равновесие между телом и излучением нарушено и тело излучает энергии больше, чем поглощает. Тогда внутренняя энергия тела начнет убывать, что приведет к уменьшению излучаемой телом энергии. Температура тела будет понижаться до тех пор, пока количество излучаемой телом энергии не станет равным количеству поглощаемой энергии. Если равновесие нарушится в другую сторону, т. е. тело начнет излучать меньше энергии, чем оно поглощает, то температура тела будет возрастать до тех пор, пока снова не установится равновесие. Таким образом, нарушение равновесия между телом и тепловым излучением вызывает процессы, направленные в сторону восстановления равновесия. Поэтому тепловое излучение называют равновесным излучением. Все другие виды излучения не являются равновесными.

Представим теперь то же самое тело, помещенное внутри другой оболочки, различающейся размерами, формой или материалом, из которого она сделана. Будем поддерживать ту же самую температуру оболочки. В системе пойдут аналогичные процессы установления равновесия, в результате которых тело внутри оболочки нагреется до той же самой температуры  $T$ . Для тела внутри оболочки ничего не изменилось: оно находится при той же самой температуре, что и прежде, и, следовательно, будет излучать ту же самую энергию. Так как тело находится в равновесии с излучением внутри оболочки, мы приходим к выводу, что *характеристики этого излучения не зависят от свойств оболочки и определяются лишь ее температурой  $T$* . Это «стандартное» излучение называется излучением **абсолютно черного тела**. Его можно охарактеризовать *плотностью энергии  $U(T)$*  — количеством энергии излучения, приходящейся на единицу объема. Очевидно, что *плотность энергии — функция температуры  $T$* .

Тепловое излучение состоит из электромагнитных волн разных частот. Полная плотность энергии складывается из плотностей энергий этих волн. Для более детальной характеристики излучения вводят дифференциальную величину — **спектральную объемную плотность энергии излучения  $u(\omega, T)$** . По определению  $u(\omega, T)$  — это энергия излучения в единице объема, приходящаяся на единичный интервал частот. Эта величина является функцией двух переменных — частоты и температуры. Иными словами, если обозначить через  $du_\omega$  энергию излучения в единице объема, приходящуюся на волны с частотами от  $\omega$  до  $\omega + d\omega$ , то  $du_\omega = u(\omega, T) d\omega$ . В СИ единицей спектральной плотности энергии является Дж·с/м<sup>3</sup>. Плотность энергии  $U(T)$  есть сумма спектральных плотностей энергии по всем возможным частотам; она выражается интегралом

$$U(T) = \int_0^{\infty} u(\omega, T) d\omega.$$

Итак, в полости существует стандартное излучение с плотностью энергии  $U(T)$ . Рассмотрим теперь тела, находящиеся с ним в равновесии.

◆ Энергия, испускаемая в единицу времени единицей поверхности тела по всем направлениям, называется **энергетической светимостью  $R(T)$**  или **интегральной плотностью потока излучения**. ◆

В СИ единицей энергетической светимости является Вт/м<sup>2</sup>. Энергетическая светимость зависит от температуры тела. Для характеристики теплового излучения важно также знать, в каком диапазоне частот излучается энергия. Поэтому вводят дифференциальную характеристику  $r_\omega = r(\omega, T)$ , называемую **испускательной способностью тела** или **спектральной плотностью потока излучения**. По определению  $r_\omega$  — это количество энергии, испускаемой в единицу времени единицей поверхности тела в единичном интервале частот. Чтобы получить энергетическую светимость тела, надо проинтегрировать испускательную способность по всем частотам:

$$R(T) = \int_0^{\infty} r_\omega(\omega, T) d\omega.$$

В СИ единицей испускательной способности является Дж/м<sup>2</sup>.

Нагретое тело не только испускает энергию, но и поглощает ее. Энергия, излучаемая всей поверхностью тела в единицу времени, называется **поток излучения**  $\Phi$ .

◆ **Поглощательная способность тела  $a_\omega$ , или спектральный коэффициент поглощения**, характеризуется отношением потока излучения, поглощенного телом в некотором малом интервале частот, к потоку падающего на тело излучения в том же интервале частот. ◆

Это безразмерная величина, не превышающая единицы.

◆ **Объект, способный поглощать при любой температуре все падающее на него излучение любых частот, называется абсолютно черным телом; для него  $a_\omega \equiv 1$ .** ◆

Будем поочередно помещать внутрь оболочки, поддерживаемой при определенной температуре  $T$ , различные тела. Очевидно, что все они находятся в одинаковых условиях, заданных излучением внутри рассматриваемой полости. Обозначим через  $f(\omega, T)$  энергию в единичном интервале частот, падающую в единицу времени на единицу поверхности любого из рассматриваемых тел в полости. Тело поглотит энергию  $a_\omega f(\omega, T)$ . В состоянии равновесия эта энергия должна быть равна испущенной телом энергии:

$$r_\omega = a_\omega f(\omega, T). \quad (33.1)$$

Различные тела в полости имеют разную поглощательную способность, следовательно, у них будет и разная испускательная способность, поэтому отношение  $r_\omega/a_\omega$  не зависит от конкретного тела, помещенного в полость:

$$\boxed{\frac{r_\omega}{a_\omega} = f(\omega, T)}. \quad (33.2)$$

В то же время испускательная способность тела зависит лишь от свойств тела, а не от полости, в которую тело помещено. Таким образом, мы приходим к заключению, что функция  $f(\omega, T)$  есть универсальная функция частоты и температуры, не зависящая ни от полости, ни от тела в ней. Эта функция, называемая функцией Кирхгофа, характеризует свойства равновесного теплового излучения.

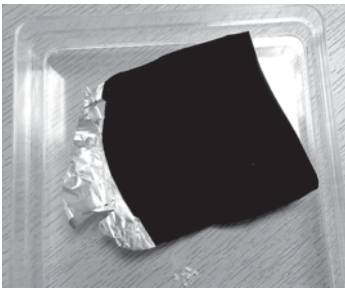
Соотношение (33.2) выражает *закон Кирхгофа*.

◆ *Отношение испускательной и поглощательной способности тела не зависит от его природы или состояния поверхности и для всех тел характеризуется универсальной функцией частоты и температуры — функцией Кирхгофа  $f(\omega, T)$ .* ◆

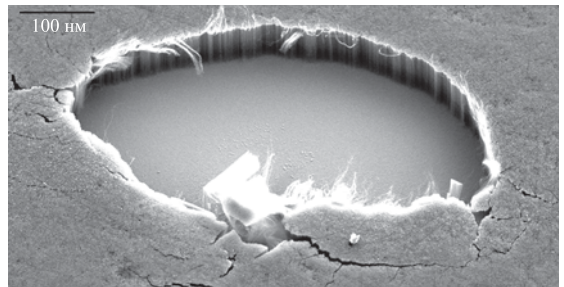
Для абсолютно черного тела (АЧТ)  $a_\omega = 1$ , откуда следует физическая интерпретация универсальной функции  $f(\omega, T)$  в законе Кирхгофа: она представляет собой испускательную способность АЧТ, т. е.  $f(\omega, T) = r(\omega, T)$ .

Почти идеальным черным телом была наша Вселенная сразу после Большого взрыва 13,7 млрд лет назад. Наполнивший Вселенную газ фотонов высоких энергий остывал по мере ее расширения, и сейчас его температура составляет всего 2,735 К, что соответствует микроволновому излучению. Это самое древнее излучение во Вселенной, его так и называют *реликтовым излучением* (другое название, также используемое в научной литературе, — *космическое микроволновое фоновое излучение*). Оно было открыто на радиотелескопе Лаборатории Белла (США) двумя инженерами — А. А. Пензиасом и Р. В. Вильсоном как ничем не устранимый фоновый «шум». Говорят, что если подсоединить телевизор к внешней антенне и переключиться на канал, не настроенный ни на одну телевизионную станцию, то около 1% мельтешения на экране приходится на долю реликтового излучения. Открытие микроволнового космического фона — яркий пример удачи в науке. Открытие было неожиданным, авторы не поняли, что именно они обнаружили, свои результаты не описали и никак не интерпретировали — это сделали уже специалисты-физики. И все равно в 1978 г. им вполне заслуженно была присуждена Нобелевская премия по физике.

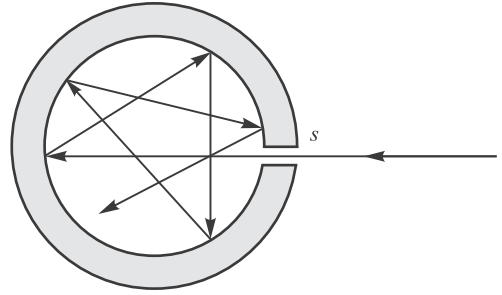
Современные нанотехнологии позволяют создать «чернейшую черноту» — защитный слой из крошечных углеродных нанотрубок, который поглощает 99,965 % падающей на него энергии. Ученые британской фирмы *Surrey Nano Systems* назвали созданный ими в 2014 г. материал *Vantablack* от англ. *black* — черный, и *vanta* — аббревиатура от слов *Vertically Aligned NanoTube Arrays* («массивы вертикально ориентированных нанотрубок»). На рисунке 33.1 (из Интернета) показана алюминиевая фольга, покрытая вантаблэком, а на рисунке 33.2 — вантаблэк под микроскопом. Этот искусственный материал чернее любой самой черной краски, изобретенной человеком. Первые эксперименты проводились со светом видимого диапазона, но в настоящее время ведутся исследования материалов, сильно поглощающих электромагнитное излучение других длин волн. Возможно, со временем «чернейшая чернота» получит практическое применение



**Рис. 33.1.** Искусственный материал *Vantablack*, выращенный на алюминиевой фольге, скрывает все детали подложки



**Рис. 33.2.** Под микроскопом *Vantablack* выглядит как густые заросли углеродных нанотрубок



**Рис. 33.3.** Полость с небольшим отверстием — модель черного тела

при изготовлении высокоточных оптических инструментов, эффективных ловителей солнечной энергии, защитных покрытий для военных объектов и т.п. Мы видим предметы благодаря отраженному от них свету. Различные детали поверхности тела различимы, так как по-разному отражают свет. На объектах, покрытых «чернейшей чернотой», невозможно различить никаких деталей, они попросту исчезают из виду, подобно черному космическому кораблю гитариста Хотблэка Дезиато<sup>1</sup>: «Он такой черный! Даже трудно определить форму — свет в него просто проваливается».

Установим теперь связь между испускательной способностью черного тела  $f(\omega, T)$  и спектральной плотностью  $u(\omega, T)$  «стандартного» излучения в полости (выше мы назвали его излучением черного тела). Сравнивая размерности этих величин, видим, что отношение функций  $f/u$  имеет размерность скорости. Единственная величина размерности скорости, которая ассоциируется с электромагнитными волнами, — это скорость света. Поэтому искомое соотношение должно иметь вид  $f(\omega, T) = \beta c u(\omega, T)$ .

Найдем безразмерный коэффициент пропорциональности  $\beta$  в этой формуле. В качестве модели черного тела возьмем замкнутую полость с небольшим отверстием площадью  $s$  (рис. 33.3). Луч света, падающий внутрь этой полости через отверстие, претерпевает многократное отражение. При каждом отражении стенки полости поглощают часть энергии. Поэтому интенсивность луча света, выходящего из отверстия, во много раз меньше интенсивности входящего луча. Чем больше отношение площади поверхности полости к площади отверстия, тем ближе такое тело к абсолютно черному. Поэтому отверстие в полости излучает как абсолютно черное тело.

В то же время внутри полости существует равновесное тепловое излучение со спектральной плотностью  $u$ . Подсчитаем энергию  $dW_\theta$ , выходящую из отверстия площадью  $s$  в направлении, заданном углом  $\theta$ . Учтем, что в данном направлении за время  $\Delta t$  может выйти только энергия, содержащаяся в наклонном цилиндре с площадью основания  $s$  и длиной образующей  $c\Delta t$  (рис. 33.4, а). Объем такого цилиндра  $V = cs\Delta t \cos \theta$ . Содержащаяся в нем энергия теплового излучения равна  $W = U(T)V = cU(T)s\Delta t \cos \theta$ . Но не вся она распространяется под углом  $\theta$ . Тепловое излучение распространяется по всем направлениям с равной вероятностью (рис. 33.4, б). Поэтому в телесный угол  $d\Omega$  попадает только часть энергии (мы обозначим ее через  $dW_\theta$ ), пропорциональная величине  $d\Omega = \sin \theta d\theta d\varphi$ . Так

<sup>1</sup> Персонаж романа «Ресторан "У конца Вселенной"» — второго из серии юмористических научно-фантастических романов Дугласа Адамса «Автостопом по Галактике».



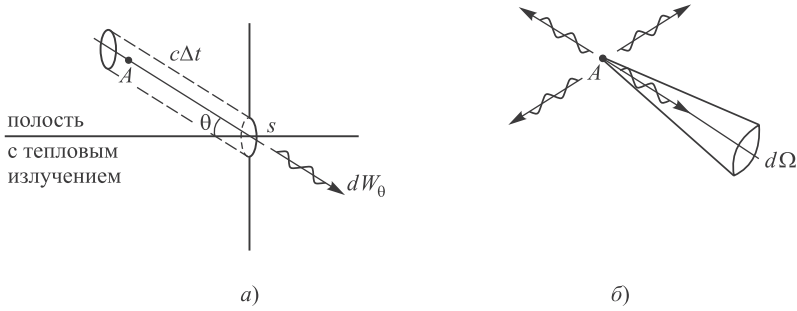


Рис. 33.4. Тепловое излучение из отверстия в полости

как полный телесный угол равен  $4\pi$ , имеем

$$dW_\theta = W \frac{d\Omega}{4\pi} = \frac{cU(T)}{4\pi} s \Delta t \sin \theta d\theta d\varphi. \quad (33.3)$$

Теперь надо проинтегрировать по углам  $\varphi$  и  $\theta$ , чтобы получить полную энергию  $\Delta W$ , выходящую из отверстия полости. Обращаем внимание: излучение падает на отверстие только из левого полупространства, так что угол  $\varphi$  может, как обычно, меняться в пределах от нуля до  $2\pi$ , но угол  $\theta$  меняется от  $0$  до  $\pi/2$ . Интегрирование по  $\varphi$  дает множитель  $2\pi$ , и нам остается взять простой интеграл по углу  $\theta$ :

$$\Delta W = \frac{cU(T)}{2} s \Delta t \int_0^{\pi/2} \sin \theta \cos \theta d\theta = \frac{cU(T)}{2} s \Delta t \frac{\sin^2 \theta}{2} \Big|_0^{\pi/2} = \frac{cU(T)}{4} s \Delta t. \quad (33.4)$$

Разделив  $\Delta W$  на время  $\Delta t$  и площадь отверстия  $s$ , получим энергетическую светимость черного тела  $R$ , а также коэффициент пропорциональности  $\beta = 1/4$ .

Итак, энергетическая светимость черного тела  $R^*(T)$  связана с плотностью энергии  $U(T)$  в полости соотношением

$$\boxed{R(T) = \frac{c}{4} U(T)}. \quad (33.5)$$

Аналогичное соотношение выполняется для дифференциальных характеристик излучения черного тела:

$$\boxed{r_\omega = f(\omega, T) = \frac{c}{4} u(\omega, T)}. \quad (33.6)$$

Таким образом, универсальная функция в законе Кирхгофа, представляющая собой испускательную способность черного тела, с точностью до множителя  $c/4$  совпадает также со спектральной плотностью равновесного теплового излучения.

До сих пор мы относили дифференциальные характеристики теплового излучения к единичному интервалу частот. Можно определить аналогичные характеристики, отнесенные к единичному интервалу длин волн. Так, если тело

испускает в интервале частот  $d\omega$  энергию  $r_\omega d\omega$ , то эту же энергию можно записать как  $r_\lambda d\lambda$ . Интервалу частот  $d\omega$  соответствует интервал длин волн  $d\lambda$ . Учитывая соотношения  $\omega = 2\pi c/\lambda$ , находим:

$$d\omega = -\frac{2\pi c}{\lambda^2}d\lambda, \quad (33.7)$$

где знак «минус» указывает на то, что с возрастанием частоты  $\omega$  длина волны  $\lambda$  убывает (в дальнейшем знак «минус» будем опускать). Таким образом:

$$r_\omega d\omega = r_\omega \frac{2\pi c}{\lambda^2}d\lambda = r_\lambda d\lambda, \quad (33.8)$$

или

$$r_\lambda = r_\omega \frac{2\pi c}{\lambda^2} = r_\omega \frac{\omega^2}{2\pi c}. \quad (33.9)$$

Аналогично связаны выражения для спектральной плотности энергии.

### 33.2 Эмпирические законы излучения абсолютно черного тела

Экспериментальное изучение распределения энергии в спектре теплового излучения АЧТ при различных температурах привело к установлению следующих закономерностей:

- спектр излучения абсолютно черного тела является сплошным, т. е. содержит непрерывный набор различных частот (длин волн);
- распределение энергии в спектре излучения неравномерно и имеет отчетливо выраженный максимум при некоторой длине волны  $\lambda_m$  (рис. 33.5);

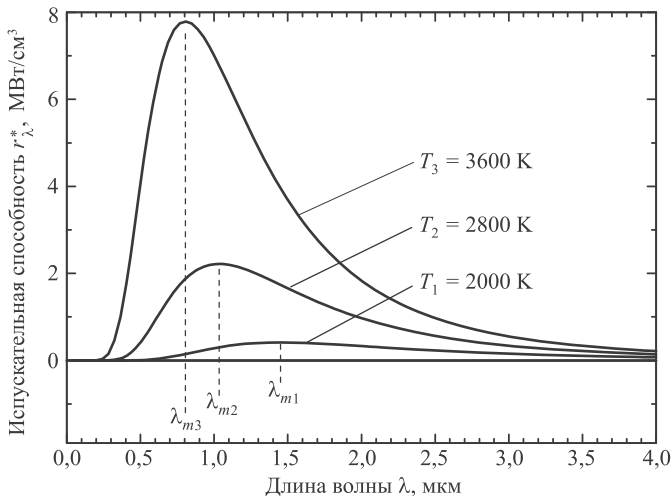


Рис. 33.5. Испускательная способность черного тела  $r_\lambda$  при разных температурах  $T$

- с повышением температуры максимум спектральной плотности излучения АЧТ смещается в сторону более коротких волн (более высоких частот) — см. рис. 33.5.

Если нагревать твердое тело, то оно сначала краснеет, а с повышением температуры свечение тела становится все более белым. Это свидетельствует о том, что максимум интенсивности теплового излучения по мере повышения температуры тела смещается к фиолетовому концу спектра, т.е. к его коротковолновой части. Длина волны  $\lambda_m$  в спектре излучения абсолютно черного тела, на которую приходится максимум спектральной плотности энергетической светимости, определяется **законом смещения Вина** (1893):

$$\lambda_m = \frac{b}{T}, \quad (33.10)$$

где **постоянная Вина**  $b = 2,90 \cdot 10^{-3}$  м·К. Закон назван по имени первооткрывателя — немецкого физика В. Вина, нобелевского лауреата (1911) «за открытия в области законов, управляющих тепловым излучением»).

Австрийские физики Й. Стефан, анализировавший экспериментальные данные (1879), и Л. Больцман, исходивший из общих термодинамических соображений (1884), нашли зависимость энергетической светимости черного тела от температуры.

♦ **Закон Стефана—Больцмана: энергетическая светимость АЧТ пропорциональна четвертой степени его термодинамической температуры:**

$$R = \int_0^{\infty} r_{\omega} d\omega = \sigma T^4. \quad \blacklozenge \quad (33.11)$$

Коэффициент пропорциональности — это так называемая **постоянная Стефана—Больцмана**  $\sigma = 5,67 \cdot 10^{-8}$  Вт·м<sup>-2</sup>·К<sup>-4</sup>.

Важно отметить, что этот закон говорит только об общей излучаемой энергии, но не о распределении энергии в спектре излучения. Постоянные Стефана—Больцмана и Вина сначала были определены эмпирически, а теоретически их значения были выведены только после создания Максом Планком квантовой теории теплового излучения, о чем речь пойдет чуть ниже.

**Задача 33.32.** Принимая, что Солнце излучает как черное тело, вычислить его энергетическую светимость и температуру поверхности. Солнечный диск виден с Земли под углом  $\theta = 32' = 9,3 \cdot 10^{-3}$  рад. Поток солнечной энергии на земной орбите (так называемая **солнечная постоянная**) равен  $C = 1,4$  кВт/м<sup>2</sup>.

**Решение.** Пусть радиус Солнца равен  $r_{\odot}$ , а расстояние до Земли равно  $l_{\oplus}$ . Отношение этих величин связано с угловым диаметром Солнца:

$$\frac{r_{\odot}}{l_{\oplus}} = \sin \frac{\theta}{2} \approx \frac{\theta}{2}. \quad (33.12)$$

Если энергетическая светимость Солнца есть  $R_{\odot}$ , то энергия, излучаемая Солнцем в единицу времени, равна произведению  $R_{\odot}$  на площадь поверхности Солнца:

$$W = R_{\odot} \cdot 4\pi r_{\odot}^2. \quad (33.13)$$

Эта энергия достигает орбиты Земли, где распределяется по большей площади  $4\pi l_{\oplus}^2$ . Отсюда находим солнечную постоянную:

$$C = \frac{W}{4\pi l_{\oplus}^2} = R_{\odot} \cdot \frac{r_{\oplus}^2}{l_{\oplus}^2} \approx R_{\odot} \frac{\theta^2}{4}. \quad (33.14)$$

В итоге получаем

$$R_{\odot} = \frac{4C}{\theta^2} = \frac{4 \times 1,4 \cdot 10^3}{(9,3 \cdot 10^{-3})^2} = 64,7 \text{ МВт/м}^2. \quad (33.15)$$

По формуле Стефана–Больцмана находим температуру верхних слоев Солнца:

$$T_{\odot} = \left( \frac{R_{\odot}}{\sigma} \right)^{1/4} = \left( \frac{64,7 \cdot 10^6}{5,67 \cdot 10^{-8}} \right)^{1/4} \approx 5800 \text{ К.} \quad \blacksquare \quad (33.16)$$

**Задача 33.33.** Исходя из данных задачи 33.32 найти длину волны, на которую приходится максимум энергии солнечного излучения.

**Решение.** Выше была определена температура верхних слоев Солнца. По закону смещения Вина получаем

$$\lambda_m = \frac{b}{T} = \frac{2,90 \cdot 10^{-3}}{5800} = 500 \text{ нм.} \quad \blacksquare$$

### 33.3 Классические результаты для теплового излучения

Несмотря на детальное изучение характеристик теплового излучения, вид обеих функций  $f(\omega, T)$  и  $u(\omega, T)$  долгое время оставался для физиков загадкой.

Попытка строгого теоретического вывода зависимости  $f(\omega, T)$  принадлежит английским ученым Дж. Рэлею и Дж. Джинсу. Воспроизведем их аргументы. Пусть полость представляет собой прямоугольный ящик с линейными размерами  $L_x, L_y, L_z$  вдоль соответствующих координатных осей. Рассмотрим стоячие электромагнитные волны вдоль оси  $x$ . Они аналогичны стоячим волнам, которые могут возбуждаться на струне. Длины стоячих волн должны удовлетворять соотношению  $L_x = n\lambda/2$ , откуда для проекции волнового вектора находим:  $k_x = \pi n/L_x$ . Целое число  $n$  нумерует разные типы волн, существующие на струне длиной  $L_x$ . Поэтому число типов стоячих волн с проекциями волнового вектора в интервале от  $k_x$  до  $k_x + dk_x$  равно:

$$dn_x = dk_x \frac{L_x}{2\pi}. \quad (33.17)$$

Мы уменьшили результат в 2 раза ввиду того, что для волнового числа здесь следует учитывать лишь абсолютные значения  $k_x$ . Действительно, волны, бегущие в противоположных направлениях и имеющие волновые числа  $+kx$  и  $-kx$ , дают одну и ту же стоячую волну. Аналогичные формулы можно записать для стоячих электромагнитных волн вдоль осей  $y$  и  $z$ :

$$dn_y = dk_y \frac{L_y}{2\pi}, \quad dn_z = dk_z \frac{L_z}{2\pi}. \quad (33.18)$$

Перемножая эти три уравнения, находим полное число типов колебаний в прямоугольной полости, волновые векторы которых лежат в интервале от  $\vec{k}$  до  $\vec{k} + d\vec{k}$ :

$$dN = 2 dn_x dn_y dn_z = 2 dk_x dk_y dk_z \frac{L_x L_y L_z}{(2\pi)^3}. \quad (33.19)$$

Здесь мы увеличили результат в 2 раза, чтобы учесть поперечность электромагнитных волн (их поляризацию): при данной длине волны световой вектор может колебаться в двух взаимно ортогональных направлениях.

Используя сферические координаты, можем записать:

$$dk_x dk_y dk_z = 4\pi k^2 dk. \quad (33.20)$$

Учитывая формулы связи частоты и длины волны  $k = 2\pi/\lambda = \omega/c$ , получаем

$$dk_x dk_y dk_z = \frac{4\pi}{c^3} \omega^2 d\omega = \frac{32\pi^4}{\lambda^4} d\lambda. \quad (33.21)$$

Произведение  $V = L_x L_y L_z$  есть объем полости. Находим тогда для плотности типов колебаний (их числа в единичном объеме полости):

$$\frac{dN}{V} = \frac{\omega^2}{\pi^2 c^3} d\omega = \frac{8\pi}{\lambda^4} d\lambda. \quad (33.22)$$

Разделив обе части уравнения (33.22) на  $d\omega$  (или  $d\lambda$ ), получим плотность числа колебаний, приходящихся на единичный интервал частот (или длин волн):

$$\frac{dN}{V d\omega} = \frac{\omega^2}{\pi^2 c^3}, \quad (33.23)$$

$$\frac{dN}{V d\lambda} = \frac{8\pi}{\lambda^4}. \quad (33.24)$$

Теперь для определения спектральной плотности излучения АЧТ  $u(\omega, T)$  и  $u(\lambda, T)$  надо умножить выражения (33.23) и (33.24) на среднюю энергию колебаний на частоте  $\omega$  или при длине волны  $\lambda$  соответственно. Классическая теорема о равномерном распределении энергии по степеням свободы гласит, что на каждую степень свободы в классической статистической системе приходится энергия  $k_B T/2$  (здесь  $k_B$  — постоянная Больцмана). У гармонического осциллятора средние кинетическая и потенциальная энергии равны между собой, и поэтому его полная средняя энергия равна  $k_B T$  (это аналогично удвоению эффективных колебательных степеней свободы молекул — см. разд. 10.4). В электромагнитной волне колеблются векторы напряженностей электрического и магнитного полей, которые дают одинаковый вклад в средний поток энергии. Если считать, что средняя энергия электромагнитной волны также равна  $k_B T$ , то для спектральной плотности энергии излучения АЧТ немедленно получаем соотношения, известные как **закон Рэлея—Джинса**:

$$\begin{aligned} u(\omega, T) &= k_B T \frac{dN}{V d\omega} = k_B T \frac{\omega^2}{\pi^2 c^3}, \\ u(\lambda, T) &= k_B T \frac{dN}{V d\lambda} = k_B T \frac{8\pi}{\lambda^4}. \end{aligned} \quad (33.25)$$

Соответственно для испускательной способности черного тела получаем:

$$\begin{aligned} r_\omega &= \frac{c}{4} u(\omega, T) = k_B T \frac{\omega^2}{4\pi^2 c^2}, \\ r_\lambda &= \frac{c}{4} u(\lambda, T) = k_B T \frac{2\pi c}{\lambda^4}. \end{aligned} \quad (33.26)$$

Для длинных волн формула Рэлея—Джинса дает хорошее совпадение с экспериментальными данными, но при коротких длинах волн, т. е. в области высоких частот, расчетные значения спектральной плотности  $r_\omega$  перестают совпадать с экспериментом (кривая закона Рэлея—Джинса уходит на бесконечность). Кроме того, энергетическая светимость черного тела также получается бесконечной:

$$R(T) = \int_0^\infty r_\omega d\omega \stackrel{?}{=} \frac{kT}{4\pi^2 c^2} \int_0^\infty \omega^2 d\omega \rightarrow \infty. \quad (33.27)$$

Создавшееся положение было названо «ультрафиолетовой катастрофой». Таким образом, классическая физика оказалась не в состоянии объяснить важные экспериментальные данные.

### 33.4 Закон излучения Планка

В 1900 г. М. Планк указал выход из создавшегося положения, выдвинув гипотезу, что *электромагнитная энергия испускается и поглощается не непрерывно, а отдельными порциями (квантами)*

$$\boxed{\varepsilon = h\nu = \hbar\omega.}$$

Коэффициенты пропорциональности  $h$  и  $\hbar = h/(2\pi)$  (оба носят название *постоянной Планка*) определены в настоящее время с большой точностью. Их современные численные значения

$$h = 6,6260755 \cdot 10^{-34} \text{ Дж} \cdot \text{с}, \quad \hbar = 1,05457266 \cdot 10^{-34} \text{ Дж} \cdot \text{с}. \quad (33.28)$$

Семь знаков после запятой! Работа экспериментаторов достойна уважения, но для наших расчетов вполне достаточно ограничиваться тремя знаками.

В соответствии с гипотезой Планка приходится исходить из того, что энергия любой колебательной системы, излучающей электромагнитные волны, может принимать лишь дискретный набор значений  $\varepsilon_n = n\hbar\omega$ , кратных собственной частоте колебаний. Отсюда Планк получил аналитическое выражение для испускательной способности черного тела.

Для излучения в состоянии равновесия по-прежнему применяется распределение Больцмана, однако гипотеза дискретности значений энергии меняет результаты самым существенным образом. Вероятность  $P_n$  того, что энергия колебания с частотой  $\omega$  равна  $\varepsilon_n = n\hbar\omega$ , определяется формулой

$$P_n = C \exp \left[ -\frac{n\hbar\omega}{k_B T} \right]. \quad (33.29)$$

Сумма всех вероятностей равна единице, откуда мы находим нормировочный коэффициент:

$$C = \frac{1}{\sum_{n=0}^{\infty} \exp[-n\hbar\omega/(k_B T)]}. \quad (33.30)$$

Средняя энергия колебаний с частотой  $\omega$  равна

$$\langle \varepsilon \rangle = \sum_{n=0}^{\infty} \varepsilon_n P_n = \hbar\omega \frac{\sum_{n=0}^{\infty} n \exp[-n\hbar\omega/(k_B T)]}{\sum_{n=0}^{\infty} \exp[-n\hbar\omega/(k_B T)]}. \quad (33.31)$$

Метод расчета таких сумм основан на выражении для геометрической прогрессии и формулы, получаемой из нее дифференцированием:

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}, \quad \sum_{n=0}^{\infty} nq^n = \frac{q}{(1-q)^2}, \quad (33.32)$$

откуда

$$\frac{\sum_{n=0}^{\infty} nq^n}{\sum_{n=0}^{\infty} q^n} = \frac{q}{1-q} = \frac{1}{1/q - 1}. \quad (33.33)$$

Подставляя сюда  $q = \exp[-\hbar\omega/(k_B T)]$ , находим выражение для средней энергии колебаний:

$$\boxed{\langle \varepsilon \rangle = \frac{\hbar\omega}{\exp[\hbar\omega/(k_B T)] - 1}}. \quad (33.34)$$

Экспоненциальное убывание средней энергии колебаний с ростом частоты решает проблему «ультрафиолетовой катастрофы»: высокие частоты вносят все меньший вклад в энергию излучения (рис. 33.6).

Подчеркнем, что главное в гипотезе Планка — это не пропорциональность энергии кванта частоте волны, а именно дискретность, «порционность» энергии излучения. Для иллюстрации предположим, что энергия кванта  $\varepsilon_x = x\omega$ , где переменная  $x$ , в отличие от теории Планка, может принимать непрерывный ряд значений. Для вероятности имеем то же самое распределение Больцмана

$$dP(x) = C \exp\left(-\frac{x\omega}{k_B T}\right) dx. \quad (33.35)$$

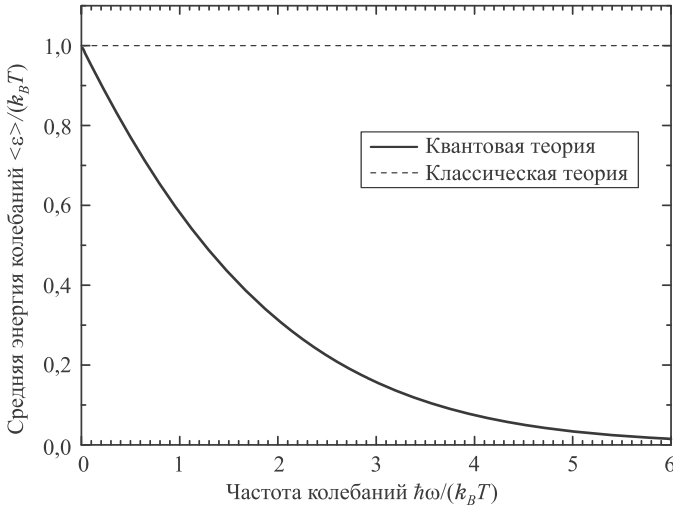
Значение константы  $C = \omega/(k_B T)$  можно найти из условия нормировки:

$$\int_0^{\infty} dP(x) = C \int_0^{\infty} \exp\left(-\frac{x\omega}{k_B T}\right) dx = C \frac{k_B T}{\omega} = 1. \quad (33.36)$$

Теперь легко получаем среднее значение энергии этой системы:

$$\langle \varepsilon \rangle = \int_0^{\infty} \varepsilon_x dP(x) = \frac{\omega}{k_B T} \int_0^{\infty} (x\omega) \exp\left(-\frac{x\omega}{k_B T}\right) dx = k_B T. \quad (33.37)$$

Таким образом, предположение о непрерывности энергии колебаний немедленно привело нас к классическому результату.



**Рис. 33.6.** Зависимость средней энергии колебаний осциллятора  $\langle \varepsilon \rangle$  от его собственной частоты в квантовой (сплошная линия) и классической (пунктирная линия) теориях; энергия выражена в «температурных единицах»  $k_B T$

Возвращаясь к результату Планка, умножаем число колебаний в единице объема, приходящихся на единичный интервал частот вблизи некоторой частоты  $\omega$  (см. уравнение (33.23)), на среднюю энергию колебаний (33.34) и получаем **формулу Планка** для спектральной плотности энергии теплового излучения абсолютно черного тела:

$$u(\omega, T) = \langle \varepsilon \rangle \frac{dN}{V d\omega} = \frac{\hbar \omega^3}{\pi^2 c^3} \cdot \frac{1}{\exp[\hbar \omega / (k_B T)] - 1}. \quad (33.38)$$

С учетом формулы (33.6) испускательная способность АЧТ описывается **законом Планка**

$$r_\omega = \frac{c}{4} u(\omega, T) = \frac{\hbar \omega^3}{4\pi^2 c^2} \cdot \frac{1}{\exp[\hbar \omega / (k_B T)] - 1}. \quad (33.39)$$

При высоких температурах (малых частотах)  $\hbar \omega = \hbar c / \lambda \ll k_B T$  и экспоненту в знаменателях (33.34) и (33.39) можно разложить в ряд:

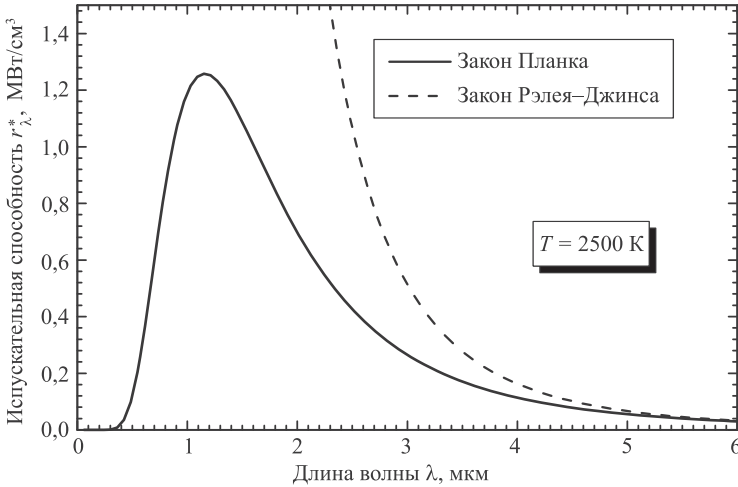
$$e^{\hbar \omega / (k_B T)} \approx 1 + \frac{\hbar \omega}{k_B T}.$$

Отсюда следует классическое выражение для средней энергии осциллятора  $\langle \varepsilon \rangle = k_B T$  и результат Рэля—Джинса (33.26).

Для спектральной плотности энергии в единичном интервале длин волн имеем выражение

$$u(\lambda, T) = \langle \varepsilon \rangle \frac{dN}{V d\lambda} = \frac{16\pi^2 \hbar c}{\lambda^5} \cdot \frac{1}{\exp[2\pi \hbar c / (\lambda k_B T)] - 1}. \quad (33.40)$$





**Рис. 33.7.** Сравнение испускательной способности  $r_\lambda$  черного тела при  $T = 2500$  К согласно закону Планка и эксперименту (сплошная линия) и закону Рэлея–Джинса (штриховая линия)

Соответствующая испускательная способность АЧТ записывается как

$$r_\lambda = \frac{c}{4} u(\lambda, T) = \frac{4\pi^2 \hbar c^2}{\lambda^5} \cdot \frac{1}{\exp[2\pi\hbar c/(\lambda k_B T)] - 1}. \quad (33.41)$$

Оказалось, что закон Планка точно согласуется с экспериментальными данными во всем спектральном диапазоне, в то время как закон Рэлея–Джинса, как уже говорилось, соответствует данным опыта только при больших длинах волн, т.е. в области, где он асимптотически совпадает с законом Планка (рис. 33.7). Более того, из закона Планка непосредственно вытекает закон Стефана–Больцмана:

$$R^*(T) = \int_0^\infty r_\omega^* d\omega = \frac{\hbar}{4\pi^2 c^2} \int_0^\infty \frac{\omega^3 d\omega}{\exp[\hbar\omega/(k_B T)] - 1}. \quad (33.42)$$

Введем безразмерную переменную интегрирования  $x = \hbar\omega/(k_B T)$ ,  $d\omega = dx k_B T/\hbar$ . В результате этого получаем

$$R(T) = T^4 \frac{k_B^4}{4\pi^2 c^2 \hbar^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} = \sigma T^4. \quad (33.43)$$

Используя выражение для интеграла

$$\int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15},$$

находим аналитическое выражение для постоянной Стефана—Больцмана:

$$\sigma = \frac{\pi^2 k_B^4}{60c^2 \hbar^3} = 5,67 \cdot 10^{-8} \frac{\text{Вт}}{\text{м}^2 \cdot \text{К}^4}, \quad (33.44)$$

величина которой согласуется с приведенными экспериментальными данными.

Из закона Планка следует также закон смещения Вина. Если продифференцировать по  $\lambda$  функцию Планка (33.41) и приравнять нулю производную  $dr_\lambda/d\lambda$ , то можно найти положение максимума функции  $r_\lambda$ . Введя безразмерную переменную  $y = 2\pi\hbar c/(k_B T \lambda)$  и обращаясь к выражению (33.41), приходим к выводу, что надо искать экстремум функции

$$\frac{1}{y^5} \frac{1}{e^y - 1},$$

дифференцирование которой приводит к трансцендентному уравнению, решаемому численно:

$$5 - y = 5e^{-y}. \quad (33.45)$$

Корень этого уравнения  $y_m = 4,965$  позволяет найти значение длины волны  $\lambda_m$ , при котором достигается максимум испускательной способности, т. е. вычислить константу Вина в законе смещения:

$$b = \lambda_m T = \frac{2\pi\hbar c}{4,965 k_B} = 2,898 \cdot 10^{-3} \text{ м} \cdot \text{К}. \quad (33.46)$$

Таким образом, формула Планка не только хорошо согласуется с экспериментальными данными, но и содержит в себе все эмпирические законы теплового излучения, а также позволяет вычислить константы в этих законах.

Мы искали максимум распределения  $r_\lambda$  по длинам волн. Но, как мы уже знаем, излучение черного тела можно характеризовать также и распределением  $r_\omega$  по частотам. Найдем для сравнения максимум этого распределения. Для этого вводим безразмерную переменную  $z = \hbar\omega/(k_B T)$  и в соответствии с (33.39) ищем экстремум функции

$$\frac{z^3}{e^z - 1}.$$

В результате дифференцирования получаем уравнение для точки максимума распределения  $r_\omega$ :

$$3 - z = 3e^{-z}, \quad (33.47)$$

которое имеет корень  $z_m = 2,821$ . Отсюда следует, что максимум интенсивности  $r_\omega$  приходится на частоту

$$\omega_m = 2,821 \frac{k_B T}{\hbar}. \quad (33.48)$$

Этой частоте соответствует длина волны

$$\lambda(\omega_m) = \frac{2\pi c}{\omega_m} = \frac{2\pi\hbar c}{2,821 k_B T}, \quad (33.49)$$

что не совпадает, естественно, с выражением для  $\lambda_m$ . Из (33.46) и (33.49) следует:

$$\frac{\lambda(\omega_m)}{\lambda_m} = \frac{4,965}{2,821} = 1,760. \quad (33.50)$$

Подчеркнем еще раз, что  $r_\lambda$  и  $r_\omega$  — разные спектральные распределения: одно из них соответствует единичному интервалу частот, другое — единичному интервалу длин волн. Эти интервалы связаны соотношением  $d\omega = 2\pi cd\lambda/\lambda^2$  и поэтому спектральные распределения  $r_\lambda$  и  $r_\omega$  не получаются друг из друга простой заменой  $\omega = 2\pi c/\lambda$ . Соответственно таким соотношением не связаны и их максимумы  $\lambda_m$  и  $\omega_m$ .

## Контрольные вопросы

1. Что такое тепловое излучение?
2. Что такое абсолютно черное тело?
3. Дайте определение спектральной плотности излучения  $u(\omega, T)$ . Какова размерность этой характеристики?
4. Дайте определение испускательной способности  $r_\omega$  тела. Какова размерность этой характеристики?
5. Дайте определение испускательной способности  $r_\lambda$  тела. Какова размерность этой характеристики?
6. Перечислите найденные на опыте законы излучения черного тела.
7. Сформулируйте закон смещения Вина и закон Стефана—Больцмана.
8. Почему спектральные распределения  $u(\omega, T)$  по частоте и  $u(\lambda, T)$  по длине волны нельзя получить одно из другого простой заменой переменных согласно соотношению  $\omega\lambda = 2\pi c$ ?
9. На каких предпосылках основан закон Рэлея—Джинса? Объясните трудности, возникшие при объяснении законов теплового излучения в рамках классической теории.
10. В чем суть гипотезы Планка об энергии излучения?
11. Сформулируйте закон Планка для спектральной плотности излучения и испускательной способности черного тела.
12. Если заглянуть в отверстие полости, стенки которой поддерживаются при определенной температуре, никакие внутренние детали (шероховатости стенок, их конфигурация и т. п.) не будут различимы. Почему?
13. В каком диапазоне электромагнитного излучения можно обнаружить черный корабль Хотблэка Дезиато (см. примечание на с. 207), если считать, что он находится при комнатной температуре  $t = 20^\circ\text{C}$ ?

## Глава 34

# ФОТОНЫ

В физике всегда так бывает: скажешь «А» в одном месте — приходится говорить «Б» в другом, даже если не очень хочется. Основная идея Планка: всякий резонатор, излучающий волны с частотой  $\omega$ , может испускать лучистую энергию лишь порциями, кратными минимальному «кванту энергии»  $\hbar\omega$ . Но физики и предположить тогда не могли, что успешное разрешение проблем с тепловым излучением приведет к ломке многих казавшихся незыблемыми представлений. И первой «пострадала» электродинамика Фарадея—Максвелла, эта «священная корова» классической физики, ее наивысшее достижение.

### 34.1 Кванты света

Благодаря Планку в физику вошла новая фундаментальная константа — постоянная Планка  $h$  или  $\hbar = h/(2\pi)$ . Эта постоянная имеет размерность момента количества движения  $L$ . В обычной жизни встречаются гораздо большие значения  $L$ , чем те, что характерны для микромира. Приведем пример медленного вращения легкого тела массой  $m$  по окружности крошечного радиуса  $R$ :

$$m = 1 \text{ г}, \quad \omega = 0,01 \text{ с}^{-1}, \quad R = 1 \text{ мм}, \quad \text{откуда}$$
$$L = m\omega R = 10^{-8} \text{ Дж} \cdot \text{с} \gg \hbar = 1,054 \cdot 10^{-34} \text{ Дж} \cdot \text{с}.$$

Пример показывает, почему в обычной жизни не наблюдается квантовая дискретность: по той же причине, по какой лестница с чрезвычайно низкими ступеньками будет восприниматься как гладкий спуск. Отсюда — способ формального перехода от квантовых результатов к классическим: надо во всех ответах устремить  $\hbar$  к нулю. При этом восстановится классическая непрерывность. Этот чисто математический прием с физической точки зрения означает, что квантовые эффекты важны для процессов, в которых константа  $\hbar$  не может считаться малой величиной.

Планк назвал величину  $h$  элементарным квантом действия. Он не питал иллюзий насчет возникающих проблем. Ведь со времен Ньютона и Лейбница, открывших дифференциальное исчисление, вся физика основывалась на непрерывности причинных соотношений. Планк отмечал поэтому, что константа  $h$

«...либо фиктивная величина, и тогда весь вывод закона излучения представлял собой всего лишь пустую игру в формулы, либо же  $h$

означает собой нечто совершенно новое и неслыханное, что должно произвести переворот в нашем физическом мышлении».

В 1905 г. А. Эйнштейн еще дальше отошел от классических представлений, предположив, что энергия не только испускается порциями, но и далее распространяется в виде индивидуальных квантов (позднее, в 1926 г., их удачно назвали **фотонами**):

«Мы должны предположить, что однородный свет состоит из зерен энергии световых квантов (*Lichtquanten*), т. е. небольших порций энергии, несущихся в пустом пространстве со скоростью света».

По Эйнштейну, энергия и импульс световых квантов связаны с соответствующими волновыми характеристиками соотношениями

$$E = h\nu = \hbar\omega, \quad p = \hbar k = \frac{h}{\lambda}. \quad (34.1)$$

Полезно представить себе классические (не квантовые) источники этих формул. В теории относительности, как известно, меняется соотношение между энергией частицы и ее импульсом: вместо хорошо знакомой формулы  $E = p^2/(2m)$  имеем

$$E = c\sqrt{m^2c^2 + p^2}. \quad (34.2)$$

Здесь  $c$  — скорость света ( $c \approx 3 \cdot 10^8$  м/с), т. е. скорость любых фотонов. Из (34.2) получаем связь скорости частицы с ее импульсом:

$$v = \frac{dE}{dp} = c \frac{p}{\sqrt{m^2c^2 + p^2}}. \quad (34.3)$$

Эта же формула вытекает из релятивистского выражения для импульса частицы (см. т. 1, разд. 9.10). Отсюда следует, что с предельной скоростью  $v = c$  могут двигаться лишь частицы нулевой массы. Полагая в (34.2)  $m = 0$ , находим для фотонов связь между энергией и импульсом:

$$E = pc. \quad (34.4)$$

В разделе 32.6 мы получили точно такое же соотношение между энергией и импульсом электромагнитной волны.

Если теперь применить к (34.4) соотношение Планка  $E = h\nu$ , то получим для импульса фотона

$$p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda},$$

поскольку  $\lambda = c/\nu$ . На основе формул (34.1) были объяснены законы фотоэффекта (см. следующий раздел).

И все-таки новые представления были весьма непривычными. Ситуация обсуждалась в 1911 г. на конгрессе с участием всех крупнейших физиков мира. Планк говорил:

«Когда думаешь о полном опытном подтверждении, которое получила электродинамика Максвелла при исследовании даже самых сложных явлений интерференции, когда думаешь о необычайных трудностях, с которыми придется столкнуться всем теориям при объяснении электрических и магнитных явлений, если они откажутся от этой электродинамики, инстинктивно испытываешь неприязнь к попыткам поколебать ее фундамент. По

этой причине мы и далее оставим в стороне гипотезу «световых квантов», тем более что эта гипотеза находится еще в зародышвом состоянии. Будем считать, что все явления, происходящие в пустоте, в точности соответствуют уравнениям Максвелла и не имеют никакого отношения к константе  $h$ ».

Итог дискуссии выразил А. Зоммерфельд (впоследствии автор первого монументального учебника по квантовой механике; среди его учеников были три лауреата Нобелевской премии по физике и два по химии):

«Я думаю, что гипотезу квантов испускания, как и начальную гипотезу квантов энергии, нужно рассматривать, скорее, как форму объяснения, а не как физическую реальность».

Итак, к 1911 г. гипотеза квантов вызывала инстинктивное ее неприятие. Но вопрос был решен экспериментаторами.

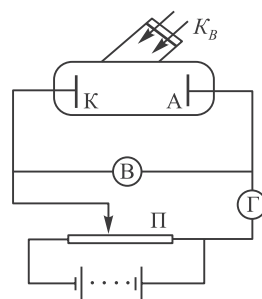
## 34.2 Фотоэлектрический эффект

Одним из явлений, подтверждающих гипотезу фотонов, является фотоэлектрический эффект (фотоэффект).

♦ *Явление вырывания электронов с поверхности твердых и жидких тел под действием света (электромагнитного излучения) носит название внешнего фотоэффекта.* ♦

Основное влияние на характеристики фотоэффекта оказывают электрические свойства облучаемого материала (проводник, полупроводник, диэлектрик), а также энергия фотонов.

Впервые явление фотоэффекта было замечено Г. Герцем в 1887 г. Первые количественные исследования внешнего фотоэффекта принадлежат русскому физику А. Г. Столетову, который опубликовал серию работ в 1888–1890 гг., и немецкому физику Ф. Ленарду (1900). В их экспериментах, схематично показанных на рис. 34.1, использовалось ультрафиолетовое излучение, энергия которого достаточно велика для вырывания электронов из металла. Свет, проникающий через кварцевое окно  $K_B$  (кварц пропускает ультрафиолетовые лучи), освещает катод  $K$ , изготовленный из исследуемого материала. Электроны, испущенные вследствие фотоэффекта, перемещаются под действием электрического поля к аноду  $A$ . В цепи возникает фототок, измеряемый гальванометром  $G$ . Потенциометром  $\Pi$  можно изменять напряжение между катодом и анодом, измеряемое вольтметром  $B$ . Исследования привели к формулировке основных законов внешнего фотоэффекта:

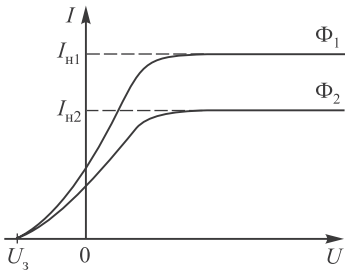


**Рис. 34.1.** Схема катодной трубки в опытах по изучению фотоэффекта

1. *Сила тока насыщения прямо пропорциональна интенсивности светового излучения, падающего на поверхность тела.*
2. *Максимальная кинетическая энергия фотоэлектронов линейно возрастает с частотой облучающего света и не зависит от его интенсивности.*

3. *Фототок не возникает, если частота света меньше некоторой определенной для данного вещества минимальной частоты (красная граница).*

Первый закон фотоэффекта был сформулирован Столетовым, красную границу фотоэффекта и рост энергии фотоэлектронов с увеличением частоты облучающего света обнаружил Ленард (Нобелевская премия в 1905 г. «за работы по катодным лучам»). Было установлено также, что фотоэффект протекает безынерционно, т. е. фототок появляется практически одновременно с освещением катода (время задержки меньше одной миллиардной доли секунды).



**Рис. 34.2.** Примерный вид вольт-амперной характеристики фототока

Проанализируем вольт-амперную характеристику (т. е. зависимость фототока  $I$  от напряжения между электродами  $U$ ), которая получается в результате фотоэлектрического эффекта. Из кривой на рис. 34.2 видно, что при некотором не очень большом напряжении фототок достигает насыщения. Это понятно даже в с точки зрения классической физики. Еще в 1898 г. Дж. Дж. Томсон установил, что частицы, выбиваемые из металла при фотоэффекте, являются электронами, открытыми им годом раньше. Насыщение фототока достигается, когда все выбиваемые электроны долетают до анода. Вместе с первым законом фотоэффекта это означает, что сила тока насыщения  $I_n$  определяется количеством электронов, испускаемых катодом в единицу времени под действием света.

Поэтому величина  $I_n$  прямо пропорциональна световому потоку:

$$I_n = k\Phi, \quad (34.5)$$

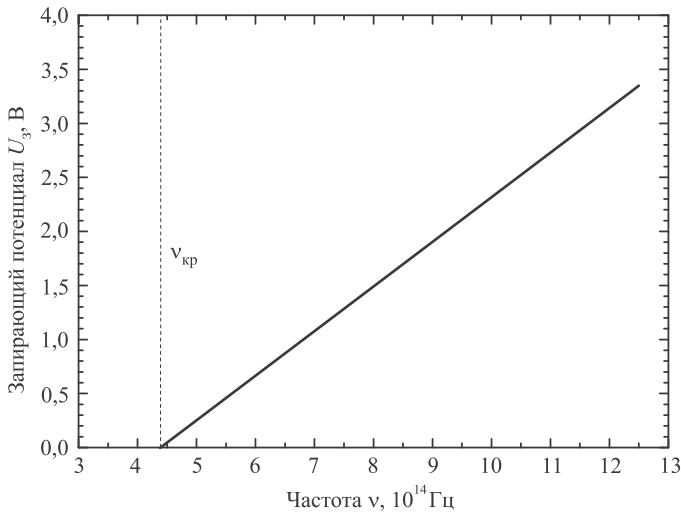
где  $k$  — коэффициент пропорциональности, характеризующий чувствительность поверхности катода к свету (его размерность  $[k] = \text{А/Вт}$ ).

Полный ход кривой показывает, что электроны вылетают из катода с различными по величине скоростями. Часть электронов, отвечающая силе тока при  $U = 0$ , обладает достаточными скоростями, чтобы долететь до анода «самостоятельно», без помощи ускоряющего поля. Для обращения фототока в нуль необходимо приложить некоторое задерживающее напряжение  $U_3$  (*запирающий потенциал*). По величине тормозящего напряжения, при котором фототок обращается в нуль, можно определить кинетическую энергию самых быстрых электронов:

$$\frac{1}{2}m_e v_{\max}^2 = eU_3, \quad (34.6)$$

где  $m_e, e, v_{\max}$  — масса, заряд и максимальная скорость электронов.

Детальные исследования фотоэффекта и измерения запирающего потенциала провел Р. Милликен, получивший Нобелевскую премию (1923) за эти исследования и измерение заряда электрона. Экспериментально было установлено, что максимальная скорость фотоэлектронов действительно зависит не от интенсивности света, а только от его частоты. Усовершенствованная аппаратура



**Рис. 34.3.** Линейная зависимость запирающего потенциала от частоты для натрия по работе Р. Милликена (*R. A. Millikan. Phys. Rev.* **7**, 355, 1916); красная граница фотоэффекта ( $\nu = 4,39 \cdot 10^{14}$  Гц,  $\lambda = 680$  нм) показана вертикальной пунктирной линией

позволила не только подтвердить рост энергии фотоэлектронов с частотой, но и доказать линейную зависимость  $U_z$  от  $\nu$  (рис. 34.3). По тангенсу угла наклона прямой к оси абсцисс  $\operatorname{tg} \alpha = h/e$  Милликен определил постоянную Планка  $h = 4,128 \cdot 10^{-15}$  эВ · с, что всего на 0,2% отличается от современного значения.

Экспериментальные результаты не укладываются в рамки классической электродинамики, так как по классическим представлениям скорость фотоэлектронов должна возрастать с увеличением амплитуды электрического поля электромагнитной волны, т.е. с ростом интенсивности облучающего света. В 1905 г. А. Эйнштейн показал, что все закономерности внешнего фотоэффекта легко объясняются, если предположить, что свет распространяется и поглощается такими же порциями (квантами)  $h\nu$ , какими он, по гипотезе Планка, испускается. Взаимодействуя с электроном вещества, фотон может обмениваться с ним энергией и импульсом. Фотоэффект возникает при неупругом столкновении фотона с электроном. При этом фотон поглощается, а его энергия передается электрону. Таким образом, электрон приобретает энергию не постепенно, а сразу — в результате единичного акта столкновения. Этим и объясняется *безынерционность фотоэффекта*.

Энергия, полученная электроном, доставляется ему в виде кванта  $h\nu$ . Часть этой энергии электрон тратит на то, чтобы вырваться из металла. Для каждого материала имеется своя *работа выхода*  $A$  — наименьшая энергия, которую необходимо сообщить электрону, чтобы удалить его из вещества в вакуум. Остаток энергии  $h\nu - A$  идет на кинетическую энергию  $E_k = m_e v^2/2$  электрона. Скорость вылетающего электрона максимальна, если он «встречается» с фотоном вблизи поверхности образца и не тратит энергию на случайные столкновения в



веществе. В этом случае будет выполняться уравнение Эйнштейна:

$$h\nu = \frac{m_e v_{\max}^2}{2} + A. \quad (34.7)$$

Нобелевская премия по физике за 1921 г. была присуждена А. Эйнштейну «за заслуги перед теоретической физикой и особенно за открытие закона фотоэлектрического эффекта». (Любопытно, что знаменитая теория относительности оказалась скрытой за общей формулировкой «за заслуги».) Уравнение Эйнштейна позволяет объяснить законы фотоэффекта. Действительно, из (34.7) непосредственно следует, что максимальная кинетическая энергия фотоэлектрона линейно возрастает с увеличением частоты подающего излучения и не зависит от его интенсивности. Так как с уменьшением частоты падающего света кинетическая энергия фотоэлектронов уменьшается (для данного вещества катода работа выхода  $A$  постоянна), при некоторой граничной частоте  $\nu = \nu_{\text{кр}}$  кинетическая энергия фотоэлектронов станет равной нулю и фотоэффект прекратится. Согласно Эйнштейну

$$\nu_{\text{кр}} = \frac{A}{h} \quad (34.8)$$

представляет красную границу фотоэффекта для данного вещества. Она зависит лишь от работы выхода электронов, т. е. от химической природы вещества и состояния его поверхности.

Используя формулу (34.6) для запирающего потенциала и выражение (34.8) для красной границы, перепишем уравнение Эйнштейна в виде

$$eU_3 = h(\nu - \nu_{\text{кр}}), \quad (34.9)$$

которое объясняет экспериментальную линейную зависимость (см. рис. 34.3) запирающего потенциала от частоты падающего электромагнитного излучения.

Таким образом, согласно Эйнштейну, свет с частотой  $\nu$  не только испускается, как это предполагал Планк, но и *распространяется* в пространстве, а также *поглощается* веществом отдельными порциями (квантами), энергия которых  $\varepsilon = h\nu = \hbar\omega$ .

«Я потратил десять лет моей жизни на проверку этого эйнштейновского уравнения 1905 года, — писал Милликен, — и вопреки всем моим ожиданиям я был вынужден безоговорочно признать в 1915 году, что оно экспериментально подтверждено, несмотря на всю свою абсурдность, поскольку, как казалось, оно противоречит всему, что мы знаем об интерференции света».

В 1914 г. были проведены модифицированные опыты по фотоэффекту:  $\gamma$ -лучи направлялись на металлическую пыль, помещенную в конденсатор. Фотоэффект практически мгновенен: при соударении пылинки с фотонами из нее выбиваются электроны, пылинка приобретает заряд и начинает двигаться в поле конденсатора. Движение пылинок наблюдалось сразу после включения источника  $\gamma$ -излучения. Если бы  $\gamma$ -излучение было классической электромагнитной волной, то волне потребовалось бы несколько секунд, чтобы раскачать электроны и вырвать их из пылинки. Отсутствие такого запаздывания воочию продемонстрировало корпускулярную природу фотоэффекта.

**Задача 34.34.** На поверхность лития падает монохроматический свет с длиной волны  $\lambda = 310$  нм. Чтобы прекратить эмиссию электронов, нужно приложить задерживающую разность потенциалов  $U_3$  не менее 1,7 В. Определить работу выхода  $A$ .

**Решение.** Энергия фотона равна  $E_{\text{ф}} = h\nu = hc/\lambda$ . Кинетическая энергия равна запирающему потенциалу  $eU_3$ . Отсюда находим работу выхода:

$$\begin{aligned} A &= E_{\text{ф}} - U_3 = \frac{hc}{\lambda} - eU_3 = \\ &= \frac{6,626 \cdot 10^{-34} \times 3 \cdot 10^8}{310 \cdot 10^{-9}} - 1,6 \cdot 10^{-19} \times 1,7 = 3,7 \cdot 10^{-19} \text{ Дж} = 2,3 \text{ эВ}. \end{aligned}$$

Внесистемная единица энергии — электронвольт ( $1 \text{ эВ} = 1,6 \cdot 10^{-19} \text{ Дж}$ ) — уже упоминалась ранее. ■

**Задача 34.35.** Уединенный цинковый шарик облучается ультрафиолетовым светом с длиной волны  $\lambda = 250 \text{ нм}$ . До какого максимального потенциала  $U_{\text{max}}$  зарядится шарик? Работа выхода электрона для цинка  $A = 3,74 \text{ эВ}$ .

**Решение.** Энергия фотонов равна:

$$E_{\gamma} = \frac{hc}{\lambda} = \frac{4,136 \cdot 10^{-15} \text{ эВ} \cdot \text{с} \times 3 \cdot 10^8 \text{ м/с}}{2,5 \cdot 10^{-7} \text{ м}} = 4,96 \text{ эВ}.$$

Выбитые электроны улетают на бесконечность, а шарик заряжается положительно. При некотором максимальном положительном потенциале шарика электроны уже не смогут преодолеть его притяжение, начнут возвращаться назад на металл и потенциал шарика перестанет возрастать. Значит, для максимального потенциала справедливо уравнение

$$eU_{\text{max}} = E_{\text{к}} = \frac{hc}{\lambda} - A = 1,22 \text{ эВ}.$$

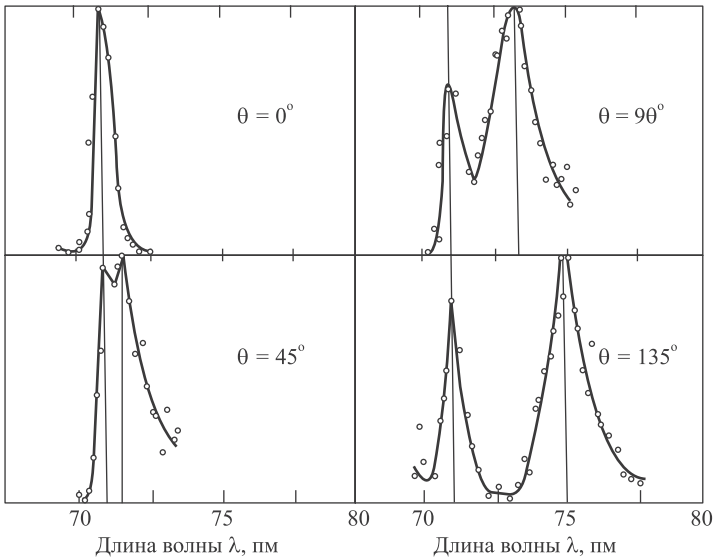
Таким образом,

$$U_{\text{max}} = \frac{E_{\text{к}}}{e} = 1,22 \text{ В}. \quad \blacksquare$$

### 34.3 Эффект Комптона

Наиболее ярко и полно корпускулярные свойства света проявляются в эффекте Комптона, заключающемся в изменении длины волны рентгеновского излучения в результате его взаимодействия с электронами вещества. Американский физик А. Комптон, исследуя рассеяние монохроматических рентгеновских лучей веществами с не очень большими атомными номерами (например бор или графит), обнаружил в 1922 г., что в составе рассеянного излучения наряду с излучением первоначальной длины волны  $\lambda$  наблюдается также излучение более длинных волн  $\lambda'$ .

С точки зрения волновой теории эффект выглядит невероятно. Если вы крикнули и голос вернулся к вам в виде эха, то он будет звучать на той же частоте. Если вы надели синий свитер и увидели себя в зеркале в красном свитере, то вы сильно удивитесь. Но именно в подобном «покраснении» рентгеновских лучей суть наблюдавшегося эффекта. Согласно классической электродинамике, электрон по действием электрического поля рентгеновского излучения должен был бы колебаться с частотой, равной частоте волны  $\nu$ . Следовательно, он должен испускать вторичное (рассеянное) излучение той же частоты и той же длины волны  $\lambda = c/\nu$ . Эксперименты А. Комптона с жестким рентгеновским излучением, рассеивающимся на электронах, показали, что это не так — длина волны рассеянного излучения увеличивалась и зависела от угла рассеяния (как



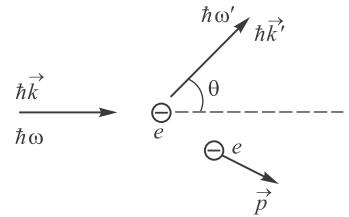
**Рис. 34.4.** Спектр гамма-излучения, рассеянного на графитовой мишени на разные углы из статьи: *A. Compton. Phys. Rev.* **22**, 409–413, 1923

это показано на рис 34.4). Открытое им явление получило простое объяснение в рамках концепции фотонов: согласно закону сохранения часть фотонного импульса передается электрону, энергия фотона уменьшается и, следовательно, уменьшается его частота — фотон «краснеет». Этот эффект незаметен при малой энергии фотонов ( $E_\gamma \ll m_e c^2$ ), т.е. для излучения с большими длинами волн, где возможно применение классической теории. В коротковолновой же области при высоких энергиях квантов ( $E_\gamma \approx m_e c^2$ ) уже проявляются корпускулярные свойства квантов излучения, необъяснимые с позиций классической теории.

Опыты показали, что разность  $\Delta\lambda = \lambda' - \lambda$  не зависит от длины волны  $\lambda$  падающего излучения и природы рассеивающего вещества, а определяется только углом рассеяния  $\theta$  между направлениями рассеянного излучения и первичного пучка:

$$\Delta\lambda = \lambda' - \lambda = 2 \frac{h}{m_e c} \sin^2 \frac{\theta}{2}, \quad (34.10)$$

где  $\lambda'$  — длина волны рассеянного излучения,  $m_e$  — масса покоя электрона,  $h$  — постоянная Планка,  $c$  — скорость света в вакууме. При рассеянии, например, на угол  $\theta = 90^\circ$  сдвиг длины волны составляет  $\Delta\lambda = h/(m_e c) = 2,4 \cdot 10^{-12} \text{ м} = 0,024 \text{ \AA}$  (здесь использована внесистемная единица длины *ангстрем*:  $1 \text{ \AA} = 10^{-10} \text{ м}$ ). В легких веществах, с которыми проводились опыты Комптона, энергия связи электрона с атомом мала по сравнению с энергией, передаваемой ему рентгеновским квантом при столкновении, т.е. в легких атомах энергией связи электрона внутри атома можно пренебречь и считать все электроны свободными. В результате своих исследований Комpton пришел к выводу, что «замечательное согласие между теорией и экспериментом ясно показывает, что это рассеяние [рентгеновских лучей] является квантовым явлением и может быть объяснено без введения каких-либо новых гипотез <...> и что кванты излучения несут как



**Рис. 34.5.** Эффект Комптона как упругое столкновение фотона с электроном

энергию, так и импульс». Формулировка Нобелевского комитета, присудившего А. Комптому эту престижную премию, была лаконичной: «За открытие эффекта, названного его именем».

Рассмотрим теперь кинематику столкновения фотона со свободным электроном, применяя релятивистские законы сохранения энергии и импульса как для соударения упругих шаров (рис. 34.5). Фотон с энергией  $\hbar\omega$  и импульсом  $\hbar\vec{k}$  ( $k = \omega/c$ ) падает на первоначально покоящийся свободный электрон ( $\vec{p}_e = 0$ ). Энергия электрона до столкновения равна  $m_e c^2$ . После столкновения электрон в результате отдачи будет обладать импульсом  $\vec{p}$  и энергией  $c\sqrt{p^2 + m_e^2 c^2}$ . Энергия и импульс фотона после рассеяния изменятся и станут равными  $\hbar\omega'$  и  $\vec{p}' = \hbar\vec{k}'$ . Запишем законы сохранения энергии и импульса:

$$\begin{aligned}\hbar\omega + m_e c^2 &= \hbar\omega' + c\sqrt{p^2 + m_e^2 c^2}, \\ \hbar\vec{k} &= \hbar\vec{k}' + \vec{p}.\end{aligned}\quad (34.11)$$

Учитывая, что  $k = \omega/c$ , перепишем закон сохранения энергии в виде

$$\sqrt{p^2 + m_e^2 c^2} = \hbar(k - k') + m_e c, \quad (34.12)$$

или

$$p^2 = \hbar^2(k^2 + k'^2 - 2kk') + 2\hbar m_e c(k - k'). \quad (34.13)$$

Уравнение закона сохранения импульса после возведения в квадрат дает

$$\begin{aligned}p^2 &= \hbar^2(\vec{k} - \vec{k}')^2 = \hbar^2(k^2 + k'^2 - 2\vec{k} \cdot \vec{k}') = \\ &= \hbar^2(k^2 + k'^2 - 2kk' \cos \theta).\end{aligned}\quad (34.14)$$

Приравнявая правые части полученных соотношений, находим:

$$m_e c(k - k') = \hbar k k' (1 - \cos \theta). \quad (34.15)$$

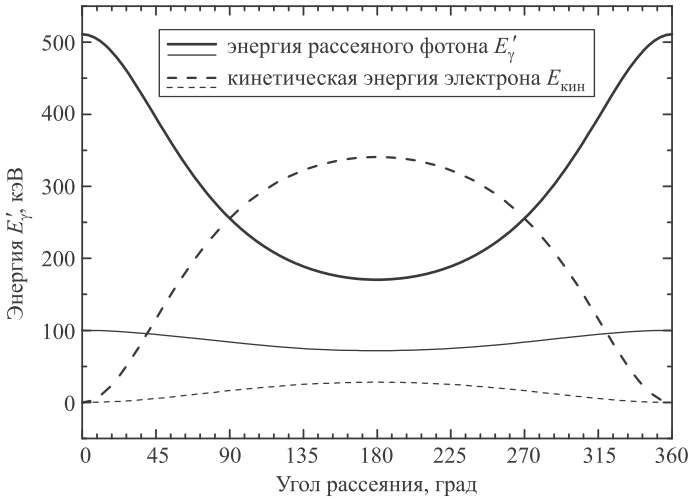
Отсюда получаем соотношение

$$\frac{2\pi}{k'} - \frac{2\pi}{k} = \frac{2\pi\hbar}{m_e c} (1 - \cos \theta). \quad (34.16)$$

Длина волны фотона связана с волновым числом ( $2\pi/k = \lambda$ ), поэтому

$$\Delta\lambda = \lambda' - \lambda = \lambda_C (1 - \cos \theta) = 2\lambda_C \sin^2 \frac{\theta}{2}, \quad (34.17)$$

где величина  $\lambda_C = 2\pi\hbar/(m_e c) = h/(m_e c)$  называется **комптоновской длиной волны** частицы с массой  $m_e$  (в данном случае электрона). Величину  $\lambda_C = \hbar/(m_e c)$  также называют комптоновской длиной волны электрона, так что мы будем их различать по обозначениям.



**Рис. 34.6.** Зависимость энергии рассеянного кванта (сплошные линии) и кинетической энергии электрона (штриховые линии) от угла рассеяния  $\theta$  фотона. Жирные линии соответствуют начальной энергии квантов, равной энергии покоя электрона ( $E_\gamma = 511$  кэВ), тонкие — энергии  $E_\gamma = 100$  кэВ

На рисунке 34.6 изображена зависимость кинетической энергии электрона и энергии кванта рассеянного излучения от угла рассеяния при двух значениях начальной энергии кванта. При энергии  $\gamma$ -кванта, равной энергии покоя электрона  $E_\gamma = 511$  кэВ, эффект Комптона явно выражен, при уменьшении же энергии до  $E_\gamma = 100$  кэВ (жесткое рентгеновское излучение) эффект уже почти незаметен.

Эксперименты показывают, что в составе рассеянного излучения присутствует несмещенная линия — излучение с начальной длиной волны (см. рис. 34.4), что можно объяснить следующим образом. Величина смещения  $\Delta\lambda$ , как мы убедились, обратно пропорциональна массе электрона. Но рентгеновское излучение может рассеиваться не только на внешних электронах атомов мишени, которые можно считать свободными, поскольку их энергия связи много меньше энергии налетающих квантов. При рассеянии части квантов на электронах внутренних оболочек атомов, связанных с ядром гораздо сильнее, энергия кванта передается атому в целом. Так как масса атома во много раз больше массы электрона, передаваемая энергия в десятки тысяч раз меньше, чем при рассеянии квантов на свободных электронах. Такой ничтожный сдвиг наблюдать невозможно, так что для таких квантов длина волны  $\lambda'$  рассеянного излучения практически не будет отличаться от длины волны  $\lambda$  падающего излучения.

**Задача 34.36.** Фотон с энергией  $E_\gamma = 2m_e c^2$  ( $m_e$  — масса электрона) при рассеянии на покоящемся электроне теряет половину своей энергии. Определить угол разлета  $\alpha$  между рассеянным фотоном и электроном отдачи.

**Решение.** В рассматриваемом случае полная энергия фотона и электрона до столкновения равна  $E = E_\gamma + m_e c^2 = 3m_e c^2$ , а после него

$$E = \sqrt{p^2 c^2 + m_e^2 c^4} + E_\gamma / 2 = \sqrt{p^2 c^2 + m_e^2 c^4} + m_e c^2,$$

где  $p$  — импульс электрона отдачи. Величина этого импульса определяется из закона сохранения энергии:

$$3m_e c^2 = \sqrt{p^2 c^2 + m_e^2 c^4} + m_e c^2.$$

Отсюда следует, что  $p = \sqrt{3}m_e c$ . А поскольку импульсы падающего и рассеянного фотонов равны  $p_\gamma = E_\gamma/c = 2m_e c$  и  $p'_\gamma = E_\gamma/(2c) = m_e c$  соответственно, имеем соотношение  $p_\gamma^2 = p^2 + p'^2_\gamma$ . Из закона сохранения импульса  $\vec{p}_\gamma = \vec{p} + \vec{p}'_\gamma$  следует тогда, что угол между импульсами электрона отдачи и рассеянного фотона  $\alpha = 90^\circ$ . ■

**Задача 34.37.** Фотон с энергией  $E_\gamma = 4m_e c^2$  рассеялся на свободном покоящемся электроне под углом  $\theta = 90^\circ$ . Определить энергию рассеянного фотона, кинетическую энергию  $E_k$  электрона и направление его движения.

**Решение.** Воспользовавшись формулой Комптона (34.10), имеем для энергии рассеянного фотона

$$E'_\gamma = \frac{hc}{\lambda'} = \frac{hc}{\lambda + \lambda_C(1 - \cos\theta)} = \frac{E_\gamma}{1 + (\lambda_C/\lambda)(1 - \cos\theta)} = \frac{E_\gamma}{1 + \frac{E_\gamma}{m_e c^2}(1 - \cos\theta)}.$$

Подставляя числа из условия задачи, находим:

$$E'_\gamma = \frac{4m_e c^2}{5} = 409 \text{ кэВ}.$$

Кинетическая энергия электрона отдачи согласно закону сохранения энергии равна разности между энергией падающего фотона и энергией рассеянного фотона:

$$E_k = 4m_e c^2 - 0,8m_e c^2 = 3,2m_e c^2 = 1,64 \text{ МэВ}.$$

Направление движения электрона отдачи найдем, применив закон сохранения импульса:

$$\vec{p}_\gamma = \vec{p}'_\gamma + \vec{p}.$$

Возводя это равенство в квадрат, получаем

$$p_\gamma^2 = p'^2_\gamma + p^2 + 2p'_\gamma p \cos \alpha,$$

где  $\alpha$  — угол между импульсами рассеянного фотона и электрона отдачи. Таким образом,

$$\cos \alpha = \frac{p_\gamma^2 - p'^2_\gamma - p^2}{2p'_\gamma p}.$$

Из закона сохранения энергии

$$E_\gamma + m_e c^2 = E'_\gamma + \sqrt{p^2 c^2 + m_e^2 c^4}$$

находим импульс электрона отдачи:

$$p^2 c^2 = (E_\gamma - E'_\gamma)^2 + 2m_e c^2 (E_\gamma - E'_\gamma).$$

Подставляя величину импульса электрона отдачи в выражение для  $\cos \alpha$ , получаем

$$\cos \alpha = \frac{(E'_\gamma - m_e c^2)(E_\gamma - E'_\gamma)}{E'_\gamma \sqrt{(E_\gamma - E'_\gamma)(E_\gamma - E'_\gamma + 2m_e c^2)}}.$$

Угол  $\alpha$  равен сумме угла  $\theta$  рассеяния фотона и угла  $\varphi$  между направлением движения падающего фотона и вектором импульса электрона отдачи:  $\alpha = \theta + \varphi$ . Так как по условию  $\theta = \pi/2$ , получаем, что  $\cos \alpha = -\sin \varphi$ . Подставляя числа, имеем

$$\sin \varphi = -\frac{(0,8 - 1)(4 - 0,8)}{0,8\sqrt{(4 - 0,8)^2 + 2(4 - 0,8)}} = \frac{0,2 \times 3,2}{0,8\sqrt{3,2 \times 5,2}} = \frac{1}{\sqrt{26}} = 0,196.$$

Следовательно,  $\varphi = 11,3^\circ$ . ■

Существует и *обратный эффект Комптона*, когда фотон, рассеиваясь на электро-не большой энергии, увеличивает свою частоту (энергию). Если провести аналогич-ные расчеты, предполагая, что электрон имеет отличную от нуля скорость  $v$ , то для длины волны рассеянного фотона получится выражение

$$\lambda' = \frac{\lambda}{1 - \frac{v}{c} \cos \alpha} \left[ 1 - \frac{v}{c} \cos \alpha' + \frac{h}{\lambda m_e c} \sqrt{1 - \frac{v^2}{c^2}} \left( 1 - \frac{v}{c} \cos \theta \right) \right], \quad (34.18)$$

где  $\alpha$  и  $\alpha'$  — углы между импульсами фотона и электрона до и после рассеяния соответственно. Для покоящегося электрона ( $v = 0$ ) отсюда следует формула (34.17).

Полную энергию  $E_e$  электрона до рассеяния можно выразить через его скорость (см. т. 1, разд. 9.12):  $E_e = m_e c^2 / \sqrt{1 - v^2/c^2}$ . Это позволяет представить множитель при последнем слагаемом в (34.18) в виде

$$\frac{h}{\lambda m_e c} \sqrt{1 - \frac{v^2}{c^2}} = \frac{hc}{\lambda E_e} = \frac{h\nu}{E_e}.$$

Следовательно, для ультрарелятивистских электронов, энергия которых намного пре-вышает энергию фотона ( $E_e \gg h\nu$ ), последним слагаемым в (34.18) можно пренебречь, и мы приходим к выражению для длины волны рассеянного фотона

$$\lambda' = \lambda \frac{1 - (v/c) \cos \alpha'}{1 - (v/c) \cos \alpha}, \quad (34.19)$$

которое можно также переписать в виде соотношения для частот:

$$\nu' = \nu \frac{1 - (v/c) \cos \alpha}{1 - (v/c) \cos \alpha'}. \quad (34.20)$$

В формулах (34.19) и (34.20) отсутствует зависимость от постоянной Планка, т. е. это уже не квантовый эффект, а классический — эффект Доплера.

Как известно, Вселенная заполнена реликтовым излучением (см. разд. 33.1). Скоп-ления (кластеры) галактик окружены облаками очень горячего ( $10^8$  К) ионизованно-го газа. В этих областях электроны, сталкивающиеся с низкочастотными фотонами реликтового излучения, увеличивают их энергию и тем самым искажают изначаль-ный спектр излучения абсолютно черного тела. Это явление называется эффектом Сюняева—Зельдовича по именам предсказавших его советских ученых (1969). Впервые наблюдавшийся в 1983 г. эффект Сюняева—Зельдовича стал мощным астрофизическим инструментом при изучении крупномасштабной структуры Вселенной.

## 34.4 Опыт Боте

Существование фотонов получило непосредственное подтверждение в опыте немецкого ученого В. Боте (1925).

Тонкая металлическая фольга F помещалась между двумя симметрично рас-положенными счетчиками  $C_1$  и  $C_2$  (рис. 34.7). Фольга освещалась слабым пучком рентгеновских лучей R, под действием которых сама становилась источником рентгеновских лучей (это явление называется рентгеновской флуоресценцией). Вследствие малой интенсивности первичного пучка число квантов, испускаемых фольгой, было невелико. Попадание рентгеновских лучей в каждый из счетчиков

вызывает немедленное срабатывание (меньше чем через  $10^{-3}$  с) и приводит в действие специальный механизм М, который делает отметку на движущейся ленте D. Если бы излучаемая энергия распространялась равномерно во все стороны, как это следует из волновых представлений, оба счетчика  $C_1$  и  $C_2$  должны были бы срабатывать одновременно и отметки на ленте приходились бы симметрично одна против другой. В действительности же наблюдалось совершенно беспорядочное расположение отметок. Это можно объяснить лишь тем, что из фольги F в отдельных актах излучения испускаются световые кванты, летящие то в одну, то в другую сторону. Таким образом, было экспериментально доказано существование особых световых частиц — фотонов, обладающих энергией и импульсом.

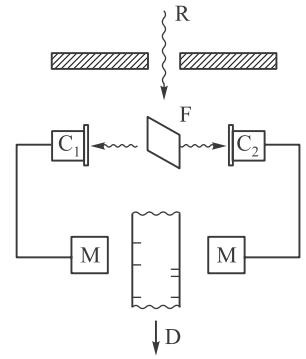


Рис. 34.7. Схема опыта Боте

Тогда же Боте разработал *метод совпадений*. В эксперименте по изучению эффекта Комптона рентгеновские лучи взаимодействовали с электронами атомов металла. Боте использовал два счетчика Гейгера, один из которых регистрировал «электроны отдачи», выбитые из атома рентгеновскими лучами, а другой — рентгеновские кванты. Совпадения происходили слишком часто, и в конечном итоге эксперимент доказал, что энергия и импульс сохраняются в каждом акте рассеяния кванта на электроне. Метод использовался в других областях физики, что принесло Боте Нобелевскую премию (1954) «за метод совпадений для обнаружения космических лучей и сделанные в связи с этим открытия».

Конкретная реализация метода совпадений зависит от исследуемой системы, но общая идея состоит в выделении событий, между которыми есть определенные корреляции. Например, открытие гравитационных волн в феврале 2016 г. основывалось на совпадении сигнала, зафиксированного практически одновременно (с разницей в 7 мс) двумя детекторами, расположенными в США на расстоянии нескольких тысяч километров друг от друга. А самое раннее (из известных авторам) использование метода совпадений относится к экспериментам фламандского математика и инженера Симона Стевина (1548–1620), который на три года раньше Галилея установил, что тяжелый и легкий шары падают на землю одновременно (1586). Эксперименты проводились на тридцатиметровой церковной башне в голландском городе Дельфте. Стевин не измерял время падения шаров по отдельности, так как такой метод заведомо не мог обеспечить достаточную точность измерений. Они с другом, философом и будущим бургомистром Дельфта Яном де Гроотом, поступили проще и эффективнее: один поднимался на башню и там одновременно выпускал из рук оба шара, а другой оставался внизу и определял на слух, есть ли интервал между звуками удара шаров о положенную под башней подставку. Разницы во времени падения обнаружено не было. (Кстати, С. Стевину мы обязаны широким распространением десятичных дробей, используемых сейчас повсеместно. Помимо этого, он ввел дробные показатели степени и пришел к закону векторного сложения сил. Стевин изобрел также двойную бухгалтерскую регистрацию — дебет/кредит.)

Подведем итоги этой главы.

- Ранее мы видели, что свет в явлениях распространения и отражения (дифракция и интерференция) ведет себя как волна с такими типично



волновыми характеристиками, как частота  $\nu$  и длина волны  $\lambda$ . Это точно установленный факт.

- Экспериментально доказано, что в явлениях испускания и передачи энергии свет ведет себя как частица, характеризующаяся энергией  $E$  и импульсом  $p$ . Это также можно считать точно установленным фактом.
- Через постоянную Планка корпускулярные характеристики численно связаны с волновыми. Поэтому приходится признать за фотоном двойственную природу. Пока в нашем курсе это необычное свойство — *корпускулярно-волновой дуализм* — установлено только для света.

## Контрольные вопросы

1. Изложите гипотезу Эйнштейна о фотонах.
2. Какова связь между волновыми характеристиками света (частота, длина волны) и корпускулярными характеристиками фотона (энергия, импульс)?
3. Энергия фотона 1 вдвое превышает энергию фотона 2. Что можно сказать о соотношении между их импульсами?
4. Покажите, что если выразить длину волны фотона  $\lambda$  в нанометрах, а энергию  $E$  — в электронвольтах ( $1 \text{ эВ} = 1,602 \cdot 10^{-19} \text{ Дж}$ ), то связь этих величин имеет вид  $E = 1240/\lambda$ . Чему равна энергия фотона ультрафиолетового излучения с  $\lambda = 100 \text{ нм}$ ? Тот же вопрос для инфракрасного излучения с  $\lambda = 10 \text{ мкм}$ .
5. Сформулируйте основные законы внешнего фотоэффекта.
6. Объясните физический смысл уравнения Эйнштейна для внешнего фотоэффекта.
7. Зависит ли энергия фотоэлектрона от интенсивности падающего на катод света?
8. Что такое красная граница внешнего фотоэффекта?
9. Определите по рисунку 34.3 работу выхода для натрия.
10. Что такое эффект Комптона? Почему эффект Комптона не наблюдается для света в видимом диапазоне?
11. В каком направлении по отношению к направлению движения падающего фотона рассеется комptonовский электрон при условии, что его кинетическая энергия имеет максимально возможное значение? В каком направлении будет при этом двигаться рассеянный фотон?
12. Почему при рассеянии рентгеновских лучей на почти свободных электронах (эффект Комптона) не происходят неупругие взаимодействия фотонов с электронами с полным поглощением фотона, как при внешнем фотоэффекте?
13. Чем объясняется появление в опытах Комптона излучения с первоначальной длиной волны?
14. В чем суть и значение опытов Боте?

## Глава 35

# Волновые свойства микрочастиц

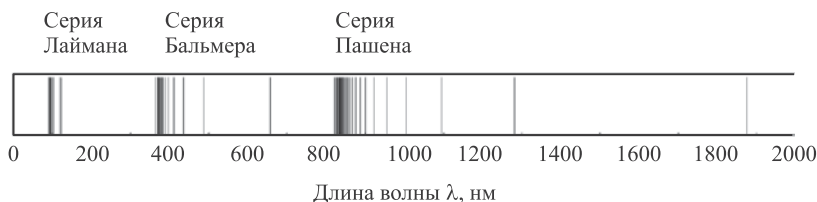
Идеи квантовой физики всякий раз радикально ломали устоявшиеся представления об устройстве мира. Мы уже видели, как на новом уровне в волновую теорию проникли, казалось, давно отброшенные наукой корпускулы — фотоны, кванты света. В этой главе мы познакомимся с тем, как волны вторглись и заняли почетнейшее место в той области физики, которая давным-давно была хорошо обжита частицами. В определенном смысле квантовая физика совершила даже более революционный переворот в наших знаниях, чем теория относительности.

### 35.1 Атом Бора

Вернемся в 1911 год. К этому времени дискретность микромира проявилась и в другом месте — в атомных спектрах. Оказалось, что атомы поглощают и испускают свет только определенной длины волны, причем спектральные линии группировались в так называемые серии (рис. 35.1).

Для спектра водорода, простейшего из атомов, была установлена (не выведена, а угадана!) несложная формула

$$\frac{1}{\lambda} = R_y \left( \frac{1}{n^2} - \frac{1}{k^2} \right), \quad R_y \approx 1,097 \cdot 10^7 \text{ м}^{-1}. \quad (35.1)$$



**Рис. 35.1.** Длины волн, излучаемых атомом водорода. Спектр состоит из серий (показаны три первые) — последовательностей линий, сгущающихся к некоторому (своему для каждой серии) предельному минимальному значению  $\lambda_{n,\min}$ . Серия Бальмера лежит в видимом диапазоне света

Здесь  $\lambda$  — длина волны излучения атома водорода,  $n$  и  $k > n$  — целые числа, а  $Ry$  — так называемая **постоянная Ридберга**. Оказалось, что серия Лаймана описывается при значениях  $n = 1, k \geq 2$ , серия Бальмера — при  $n = 2, k \geq 3$ , а серия Пашена — при  $n = 3, k \geq 4$ . Предельные минимальные значения для длин волн в каждой серии получаются при  $k \rightarrow \infty$ :  $\lambda_{n,\min} = 91,16n^2$  нм.

Кроме того, из свойств газов давно было известно, что размеры атомов приблизительно равны  $10^{-10}$  м. Требовалась теория для объяснения атомных спектров и размеров атомов. Она должна была включать в себя некий параметр, позволяющий построить величины, имеющие размерность длины (постоянных  $e$  и  $m_e$  для этого недостаточно). Такого параметра в классической теории не было. Им могла бы стать постоянная Ридберга, но ее происхождение было темно и загадочно.

Опыты Резерфорда в 1911 г. по отклонению  $\alpha$ -частиц при падении на металлическую фольгу показали наличие внутри атома положительного заряда, сконцентрированного в области диаметром не более  $10^{-14}$  м. Возникла планетарная модель атома водорода: протон с электроном на орбите. Физики любят единые модели, а здесь так красиво в малом повторялось большое, в атоме — Солнечная система.

Проблема состояла в том, что электрон должен упасть на ядро. Почему планеты не падают на Солнце? Важны начальные условия (положение и скорость), а дальше под действием силы тяготения планета вращается по стационарной орбите, т. е. движется с центростремительным ускорением. Электрон, в отличие от планеты, заряжен, а по законам классической электродинамики заряженная частица при ускорении испускает электромагнитное излучение (см. разд. 32.10). Потеряв свою энергию, электрон должен бы быстро упасть на ядро, так что свет за это время даже не успел бы достичь наших глаз. В таком случае получалось бы, что мы видим лишь уже «мертвые» атомы, а это не так! Резерфорд понимал проблему, но сознательно сконцентрировался на изучении ядра, полагая, что вопрос об устойчивости атома будет решен при исследовании поведения атомных электронов. Это суждено было сделать в 1913 г. Нильсу Бору, предложившему новую теорию атома.

## Постулаты Бора

◆ *В атоме существуют стационарные орбиты, находясь на которых электрон не излучает.* ◆

Здесь прослеживается «насильственное» введение дискретности (разрешены не все орбиты), а также типичное для физики «заметание проблемы под ковер»: если чему-то не находится объяснений, принимают это как данность и изучают следствия в надежде, что когда-нибудь поймут и причину.

◆ *При переходе с одной стационарной орбиты на другую электрон излучает (или поглощает) квант электромагнитного излучения с частотой  $\omega = \Delta E/\hbar$ , где  $\Delta E$  — разность энергий электрона на этих орбитах.* ◆

Этот постулат отражает сохранение энергии и соотношения Планка—Эйнштейна.

◆ *Динамика электрона на стационарной орбите определяется уравнениями классической теории.* ◆

Неизбежное следствие: так как остальные орбиты для электрона запрещены, переход осуществляется скачком; о пути и энергии электрона между орбитами говорить не имеет смысла: законы механики там неприменимы.

◆ *Круговые стационарные орбиты определяются условием квантования момента количества движения ( $n$  — целое число):*

$$L = n\hbar. \quad \blacklozenge$$

Постоянная Планка  $\hbar$  имеет размерность момента количества движения и вместе с зарядом электрона  $e$  и его массой  $m_e$  позволяет образовать параметр, имеющий размерность длины. Это приводит к возможности вычислить размеры атома.

Согласно классической механике, в водородоподобном атоме электрон может вращаться по орбите радиусом  $R$  со скоростью  $v$  вокруг ядра с зарядом  $Ze$  ( $Z$  — целое число, атомный номер ядра), причем эти величины связаны между собой вторым законом Ньютона:

$$\frac{m_e v^2}{r} = \frac{Ze^2}{4\pi\epsilon_0 r^2}, \quad \text{откуда} \quad v = \frac{e\sqrt{Z}}{\sqrt{4\pi\epsilon_0 m_e r}}. \quad (35.2)$$

Поэтому энергия и момент импульса электрона в таком атоме выражаются через радиус орбиты как

$$\begin{aligned} E &= \frac{mv^2}{2} - \frac{Ze^2}{4\pi\epsilon_0 r} = -\frac{Ze^2}{8\pi\epsilon_0 r}, \\ L &= m_e v r = e\sqrt{\frac{Zm_e r}{4\pi\epsilon_0}}. \end{aligned} \quad (35.3)$$

Если к последнему уравнению применить *условие квантования Бора*  $L = n\hbar$  ( $n = 1, 2, 3, \dots$ ), то получатся выражения для энергии электрона на стационарной орбите, радиусов орбит и прочих характеристик водородоподобного атома в модели Бора.

## Характеристики водородоподобного атома

- Радиусы разрешенных орбит

$$r_n = a_B \frac{n^2}{Z}, \quad a_B = \frac{4\pi\epsilon_0 \hbar^2}{m_e e^2}. \quad (35.4)$$

- Энергия электрона на стационарной орбите

$$E_n = -\frac{Z^2 e^2}{8\pi\epsilon_0 a_B n^2} = -\frac{Z^2 m_e e^4}{2(4\pi\epsilon_0)^2 \hbar^2 n^2} = -13,6 \text{ эВ} \cdot \frac{Z^2}{n^2}. \quad (35.5)$$

Константа  $a_B$ , имеющая размерность длины, называется *радиусом Бора* ( $a_B = 5,29 \cdot 10^{-11}$  м). Смысл числа  $n$  — номер разрешенной орбиты. Радиус Бора — радиус низшей орбиты ( $n = 1$ ) в атоме водорода ( $Z = 1$ ). При  $h \rightarrow 0$  получаем, что радиус Бора стремится к нулю: падение электрона на ядро, как в классике!

Формула (35.5) определяет разрешенные значения энергии в атоме водорода, или, как говорят, энергетические уровни. Отрицательные  $E$  соответствуют связанным состояниям с движением в ограниченной области пространства (аналог в классической физике — движение планет по эллипсам в отличие от гиперболических траекторий, уходящих на бесконечность).

При решении задач о поведении электрона в атоме обычно возникает квадрат электрического заряда электрона в комбинации с электрической постоянной  $\epsilon_0$ . Для выполнения численных расчетов удобно выразить эту комбинацию через так называемую *постоянную тонкой структуры*, обычно обозначаемую греческой буквой «альфа»:

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{137,036}. \quad (35.6)$$

Постоянная тонкой структуры, введенная в физику А. Зоммерфельдом, является фундаментальной физической постоянной, характеризующей силу электромагнитного взаимодействия, и играет важную роль в квантовой электродинамике. Приведем для справки формулу для энергии (35.5), выраженной через постоянную тонкой структуры:

$$E_n = -\alpha^2 m_e c^2 \frac{Z^2}{2n^2}. \quad (35.7)$$

Из-за множителя  $\alpha^2$  характерные для атома энергии на четыре порядка меньше энергии покоя электрона. Это — проявление «нерелятивизма» атомной физики.

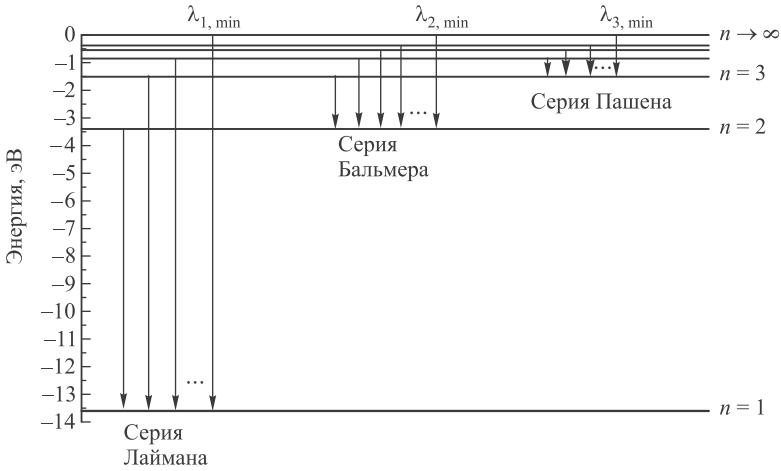
При переходе с уровня  $k$  на уровень  $n$  ( $k > n$ ) излишек энергии перейдет в энергию фотона:  $E_k - E_n = h\nu_{nk}$ . Поэтому для спектра излучаемых длин волн  $\lambda_{nk} = c/\nu_{nk}$  получаем из (35.5) соотношение

$$\frac{1}{\lambda_{nk}} = R_y Z^2 \left( \frac{1}{n^2} - \frac{1}{k^2} \right), \quad (35.8)$$

$$R_y = \frac{m_e e^4}{8c\epsilon_0^2 h^3} = 1,097 \cdot 10^7 \text{ м}^{-1},$$

совпадающее при  $Z = 1$  с эмпирической формулой Ридберга (35.1). Таким образом, теория Бора позволила также вычислить постоянную Ридберга  $R_y$ . Стали понятны и существование спектральных серий, и предельные значения  $\lambda_{n,\min}$  (рис. 35.2). В 1922 г. Нильс Бор получил Нобелевскую премию «за заслуги в исследовании строения атомов и испускаемого ими излучения».

При выводе формулы (35.5) для энергетических уровней водородоподобного иона был сделан ряд допущений.



**Рис. 35.2.** Схема энергетических уровней атома водорода по модели Бора

Оценим возникающие из-за этого погрешности. Во-первых, кулоновский центр предполагался неподвижным, т. е. масса ядра считалась бесконечно большой. На самом деле масса ядра  $M$  конечна, и оно вместе с электроном вращается вокруг общего центра масс. Эта задача, как известно из классической механики, эквивалентна задаче о вращении частицы с приведенной массой  $\mu = m_e M / (m_e + M)$ . Именно эта масса должна стоять в (35.5) вместо массы электрона. Так как  $M \gg m_e$ , то  $\mu \approx m_e(1 - m_e/M)$ . Энергия уровня линейно зависит от массы электрона, и, значит, поправка  $\Delta E$  к энергии за счет конечности массы ядра составляет для водорода  $\Delta E \approx -m_e/m_p = 0,5 \cdot 10^{-3}$  и приводит к уменьшению энергии связи, т. е. к повышению уровня. Для основного уровня атома водорода это составляет всего  $0,007 \text{ эВ} = 7 \text{ мэВ}$  (около  $0,05\%$ ).

Релятивистские поправки связаны с использованием нами классического выражения  $T_{\text{кл}} = m_e v^2 / 2$  для кинетической энергии электрона. Релятивистскую формулу (см. т. 1, разд. 9.11) мы разложим в ряд по степеням отношения скорости электрона к скорости света:

$$T = m_e c^2 \left( \frac{1}{\sqrt{1 - v^2/c^2}} - 1 \right) = m_e c^2 \left( 1 + \frac{1}{2} \frac{v^2}{c^2} + \frac{3}{8} \frac{v^4}{c^4} + \dots - 1 \right) \approx \frac{m_e v^2}{2} \left( 1 + \frac{3}{4} \frac{v^2}{c^2} \right).$$

Отсюда относительная поправка к энергии уровня

$$\frac{\Delta E}{E} \approx \frac{v^2}{c^2} \approx \frac{E}{m_e c^2} \approx 10^{-5},$$

что приводит к увеличению энергии связи (понижению уровня) примерно на  $0,001\%$ .

При выводе формулы (35.5) предполагалась также точность кулоновского центра притяжения. На самом деле атомное ядро (и даже протон) имеет конечные размеры. Согласно квантовой механике, имеется отличная от нуля вероятность, что электрон окажется внутри ядра, а там (при сферическом распределении электрического заряда) кулоновское притяжение слабее и, стало быть, меньше энергия связи — уровень повышается. Относительная величина поправки оказывается порядка квадрата отношения радиуса ядра к радиусу Бора:  $\Delta E/E \approx (R/a_B)^2$ . Радиус протона  $R \approx 10^{-15} \text{ м}$ , так что

для водорода относительная поправка на размер ядра составляет  $\Delta E/E \approx 10^{-10}$ . Эта поправка ничтожно мала.

Еще имеются поправки, связанные с наличием у электрона и протона магнитных моментов (это приводит к расщеплению уровней энергии, или к тонкой структуре уровней), а также с эффектами, возникающими при квантовании электромагнитного поля. Но о них на момент создания модели атома Бора вообще даже не подозревали.

**Задача 35.38.** Определить скорость электрона на  $n$ -й орбите атома Бора.

**Решение.** Радиус  $n$ -й орбиты определяется формулой  $R_n = a_B n^2$ , где  $a_B$  — радиус Бора. Скорость электрона  $v$  выразим через момент количества движения  $L = n\hbar$ :

$$v_n = \frac{n\hbar}{m_e R_n} = \frac{\hbar}{m_e a_B n}.$$

Выражение для радиуса Бора можно упростить, используя введенную постоянную тонкой структуры:

$$a_B = \frac{4\pi\epsilon_0 \hbar^2}{m_e e^2} = \frac{4\pi\epsilon_0 \hbar c}{e^2} \frac{\hbar}{m_e c} = \frac{1}{\alpha} \frac{\hbar}{m_e c}. \quad (35.9)$$

Подставляя это выражение в полученную выше формулу для скорости электрона, получаем для  $n$ -й орбиты

$$v_n = \frac{\alpha c}{n}. \quad (35.10)$$

Отсюда вытекает, что на низшей орбите скорость электрона приблизительно в 137 раз меньше скорости света ( $v_1 = 2,2 \cdot 10^6$  м/с), т.е. атом — это действительно нерелятивистская система. На  $n$ -й орбите скорость электрона в  $n$  раз меньше, чем на первой.

**Задача 35.39.** На основе законов классической механики и электродинамики оценить время жизни  $t_{ж}$  электрона на первой орбите водородоподобного атома Бора.

**Решение.** При вращении электрон должен терять энергию и опускаться на ядро по спирали. Для оценки времени жизни электрона предположим, что в каждый данный момент времени, когда он находится на расстоянии  $r$  от ядра, его траекторию можно аппроксимировать окружностью радиусом  $r$ . В соответствии с уравнением (35.2) находим тогда центростремительное ускорение электрона в этой точке орбиты:

$$a = \frac{v^2}{r} = \frac{Z e^2}{4\pi\epsilon_0 m_e r^2}. \quad (35.11)$$

Подставляя это ускорение в формулу Лармора (32.137), находим мощность электромагнитного излучения:

$$N = \frac{2}{3} \frac{e^2 a^2}{4\pi\epsilon_0 c^3} = \frac{2}{3} Z^2 \left( \frac{e^2}{4\pi\epsilon_0 c} \right)^3 \frac{1}{m_e^2 r^4}. \quad (35.12)$$

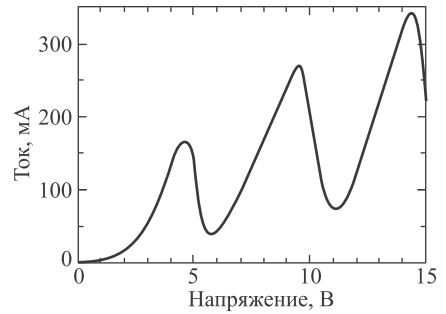
Зависимость энергии электрона от расстояния до ядра выражается формулой (35.3). Дифференцируем ее по времени:

$$\frac{dE}{dt} = \frac{Z}{2} \left( \frac{e^2}{4\pi\epsilon_0} \right) \frac{1}{r^2} \frac{dr}{dt}. \quad (35.13)$$

По закону сохранения энергии мощность излучения должна быть равна потере энергии электроном в единицу времени:  $-dE/dt = N$ . Используя (35.12) и (35.13) и разделяя переменные, получаем уравнение

$$-r^2 dr = \frac{4}{3} Z \left( \frac{e^2}{4\pi\epsilon_0 c} \right)^2 \frac{1}{m_e^2 c} dt. \quad (35.14)$$

**Рис. 35.3.** Зависимость анодного тока от приложенного напряжения в опытах с парами ртути (график, взятый из Википедии, основан на результатах оригинальной работы Франка—Герца); периодичность пиков составляет 4,9 В



Это уравнение надо проинтегрировать по времени от 0 до  $t_{\text{ж}}$ , а по расстоянию  $r$  от радиуса первой орбиты  $r_1 = a_B/Z$  (см. (35.4)) до нуля:

$$-\int_{r_1}^0 r^2 dr = \frac{a_B^3}{3Z^3} = \frac{4}{3}Z \left( \frac{e^2}{4\pi\epsilon_0 c} \right)^2 \frac{t_{\text{ж}}}{m_e^2 c}. \quad (35.15)$$

Подставляя сюда выражение для радиуса Бора, получаем искомую оценку:

$$t_{\text{ж}} = \frac{1}{4Z^4 \alpha^5} \frac{\hbar}{m_e c^2}. \quad (35.16)$$

Для упрощения формул и облегчения численных расчетов мы использовали выражение (35.6) для постоянной тонкой структуры.

Подставляя в (35.16) численные значения  $\hbar = 6,58 \cdot 10^{-16}$  эВ·с,  $m_e c^2 = 0,511 \cdot 10^6$  эВ и  $\alpha = 1/137$ , находим:

$$t_{\text{ж}} = \frac{137^5}{4Z^4} \times \frac{6,58 \cdot 10^{-16}}{0,511 \cdot 10^6} \approx \frac{1,6 \cdot 10^{-11}}{Z^4} \text{ с}. \quad (35.17)$$

Заметим сильную зависимость времени жизни «классического» электрона от заряда ядра  $Z$ . Но даже для атома водорода ( $Z = 1$ ) время  $t_{\text{ж}}$  мало, за это время свет успевает пройти расстояние  $ct_{\text{ж}} \approx 5$  мм. ■

Гипотеза Бора о существовании дискретных энергетических уровней атома нашла экспериментальное подтверждение в опытах немецких физиков Дж. Франка и Г. Герца (племянника знаменитого Генриха Герца, первооткрывателя электромагнитных волн). Они изучали вольт-амперную характеристику при бомбардировке паров ртути катодными электронами в вакуумной трубке. При напряжении ускоряющего потенциала 4,9 В и кратных значениях 9,8 В, 14,7 В и т. д. наблюдался резкий спад анодного тока (рис. 35.3). Одновременно наблюдалось ультрафиолетовое свечение паров ртути на длине волны  $\lambda = 253,7$  нм, что соответствует энергии фотона  $\hbar\omega = 4,9$  эВ. Причина такого поведения вольт-амперной характеристики состоит в соударениях вылетающих из катода электронов с атомами ртути. При небольшой величине ускоряющего потенциала электроны сталкиваются с атомом ртути упруго, их энергия не меняется и ток растет монотонно с ростом напряжения на сетке вакуумной трубки. Когда ускоряющий потенциал достигает значения 4,9 В, энергия электронов целиком идет на возбуждение атомов. Полностью потерявшие свою энергию электроны останавливаются, и анодный ток резко падает. А возбужденные атомы ртути



переходят потом в основное состояние, излучая фотон соответствующей энергии. При дальнейшем увеличении ускоряющего потенциала у электронов после столкновения с атомами еще остается часть кинетической энергии, и потому ток на аноде увеличивается с ростом напряжения. По достижении значения потенциала 9,8 В у электронов уже хватает энергии на последовательное возбуждение двух атомов ртути — анодный ток снова резко падает. Чередование максимумов тока при значениях напряжения, кратных величине 4,9 В, соответствующей энергии излучения возбужденных атомов ртути, доказало дискретность энергетических состояний атома. В 1925 г. Дж. Франк и Г. Герц были удостоены Нобелевской премии «за открытие законов соударения электрона с атомом».

**Задача 35.40.** При переходе с третьего уровня на второй (головная линия серии Бальмера) водородоподобный ион атома некоторого элемента испускает фотон с энергией  $W = 7,5$  эВ. Какой это элемент?

**Решение.** Энергия электрона, находящегося на  $n$ -й орбите вокруг ядра с зарядом  $Ze$ , равна:

$$E_n = -13,6 \text{ эВ} \cdot \frac{Z^2}{n^2}.$$

При переходе с уровня  $n = 3$  на уровень  $n = 2$  выделяется энергия

$$W = E_3 - E_2 = 13,6 \text{ эВ} \cdot Z^2 \left( \frac{1}{4} - \frac{1}{9} \right) = 13,6 \text{ эВ} \cdot \frac{5}{36} Z^2,$$

откуда

$$Z = \sqrt{\frac{36 \cdot 7,5}{5 \cdot 13,6}} = 1,99.$$

Атомный номер элемента — целое число, так что полученный результат следует округлить и считать  $Z = 2$ , что соответствует гелию. ■

Еще до появления теории Бора, как мы уже знаем, был изучен спектр водородного атома и эмпирически установлена формула (35.1). Но при наблюдении спектра Солнца были замечены линии, казалось бы нарушающие эту формулу, так как они соответствовали полуцелым значениям  $n$  и  $k$ . После появления теории Бора стало ясно, что квантовые числа  $n$  и  $k$  все-таки должны быть целыми, а кажущиеся полуцелые значения можно объяснить по-другому. Действительно, из формулы (35.8) для частот, испускаемых водородоподобным атомом, следует, что

$$\nu_{nk} = cRy \left( \frac{1}{(n/2)^2} - \frac{1}{(k/2)^2} \right) = cRy(2^2) \left( \frac{1}{n^2} - \frac{1}{k^2} \right),$$

т. е. наблюдавшиеся линии принадлежат иону элемента с  $Z = 2$ . Как известно, этот элемент носит «солнечное» имя — гелий.

**Задача 35.41.** Определить скорость  $v$ , которую приобретает атом водорода в результате излучения кванта света при переходе электрона со второго уровня на первый. Найти, как увеличится в результате отдачи длина волны излучаемого кванта?

**Решение.** Массу атома водорода можно принять равной массе протона  $m_p$ . Кинетическая энергия атома в результате отдачи приобретает за счет уменьшения энергии (частоты) испускаемого кванта:

$$\frac{m_p v^2}{2} = h\Delta\nu. \quad (35.18)$$

Импульс испущенного кванта должен быть равен импульсу, приобретенному атомом:

$$m_p v = \frac{h}{c}(\nu - \Delta\nu), \quad (35.19)$$

где  $h\nu = \Delta E$  — разность энергий второго и первого уровней (иными словами,  $\nu$  — частота фотона, испускаемого в отсутствие отдачи атома). Из-за большой массы атома  $\Delta\nu/\nu \ll 1$ , тогда сдвигом частоты в (35.19) можно пренебречь и сразу найти скорость атома:

$$v \approx \frac{h\nu}{m_p c} = c \frac{\Delta E}{m_p c^2}. \quad (35.20)$$

Подставляя (35.20) в (35.18), находим относительное изменение частоты испущенного фотона:

$$\frac{\Delta\nu}{\nu} = \frac{h\nu}{2m_p c^2} = \frac{\Delta E}{2m_p c^2}. \quad (35.21)$$

Разность энергий второго и первого уровней атома водорода

$$\Delta E = 13,6 \text{ эВ} (1 - 1/4) = 10,2 \text{ эВ},$$

энергия покоя протона  $m_p c^2 = 938,3 \text{ МэВ}$ . Отсюда  $\Delta\nu/\nu = 0,543 \cdot 10^{-8}$ . Для скорости отдачи получаем тогда из (35.20)

$$v = 2 \frac{\Delta\nu}{\nu} = 2 \times 3 \cdot 10^8 \times 0,543 \cdot 10^{-8} = 3,26 \text{ м/с}.$$

Учитывая, что  $\Delta\lambda/\lambda = \Delta\nu/\nu$ , и находя из (35.1) обратную длину волны первой линии серии Лаймана

$$\frac{1}{\lambda} = \frac{3}{4} R_y = 0,75 \times 1,097 \cdot 10^7 \text{ м}^{-1} = 0,823 \cdot 10^7 \text{ м}^{-1},$$

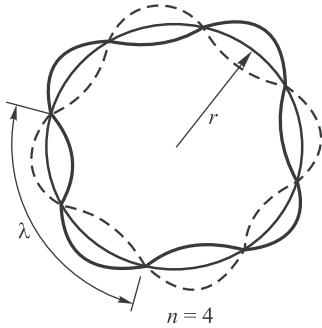
получаем искомое изменение длины волны

$$\Delta\lambda = \frac{0,543 \cdot 10^{-8}}{0,823 \cdot 10^7} = 6,6 \cdot 10^{-16} \text{ м. } \blacksquare$$

## 35.2 Волны материи

Итак, в модели атома Бора снова появилась дискретность, возникли загадочные целые числа. В классической физике они появлялись в явлениях интерференции (номера максимумов и минимумов) и в стоячих волнах (число узлов на закрепленных струнах). Французский физик Луи де Бройль выдвинул фундаментальную гипотезу (1923): **♦ корпускулярно-волновой дуализм присущ всем микрочастицам. ♦**

Де Бройль предложил ассоциировать с каждой микрочастицей некую волну. Тогда условие стационарности орбиты в атоме Бора можно сформулировать как требование, чтобы электронная волна была стоячей. В этом случае узлы волны неподвижны, а пучности приходятся на одни и те же точки электронной орбиты — очень наглядная интерпретация стационарных орбит в модели Бора. Но для этого длина волны должна укладываться на длине окружности орбиты целое число раз, и квантование орбит возникает естественным путем: если  $\lambda$  — длина волны электрона на орбите радиусом  $r$ , то для стоячей волны верно



**Рис. 35.4.** Электронная орбита и стоячая волна де Бройля с четырьмя узлами

соотношение  $2\pi r = n\lambda$  (рис. 35.4). В качестве аналогии можно привести пример с обклеиванием комнаты обоями: если по периметру комнаты укладывается целое число периодов рисунка на обоях, то при дальнейшем обклеивании рисунок воспроизводится на тех же местах — своего рода стационарность.

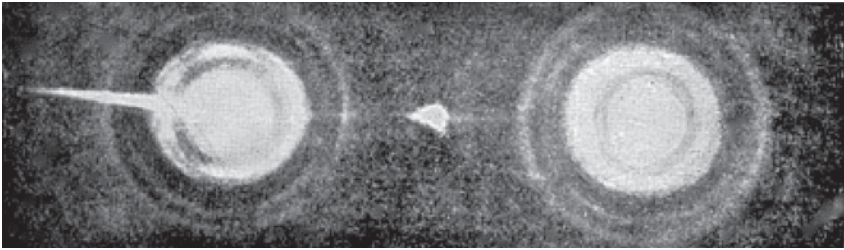
Условие квантования Бора  $L = pr = n\hbar$  вытекает теперь из условия де Бройля для стоячей волны, если длину волны де Бройля связать с импульсом электрона:

$$\lambda = \frac{h}{p}. \quad (35.22)$$

Точно такое же соотношение (34.1) связывает импульс и длину волны фотона. Долгое время уделялось внимание только волновым свойствам фотона, но в самом начале XX в. восстановили в правах корпускулярные. С электроном получилось наоборот: все видели в нем частицу, и лишь де Бройль разглядел в нем волну. В 1929 г. Луи де Бройлю была присуждена Нобелевская премия «за открытие волновой природы электронов». Более того, формула (35.22) верна не только для фотонов и электронов, но и для любых других частиц.

Гипотеза де Бройля подтверждена экспериментами К. Дэвиссона и Л. Джермера из Лаборатории Белла в США (отражение электронов от кристаллических плоскостей) и Дж. П. Томсона из Лондонского университета (рассеяние электронов на фольге), обнаруживших (1927) на фотопластинках типичные для оптики дифракционные кольца (рис. 35.5). В 1937 г. Дэвиссон и Томсон получили Нобелевскую премию «за экспериментальное открытие дифракции электронов на кристаллах». Кстати, Джордж Паджет Томсон — сын знаменитого Джозефа Джона Томсона, первооткрывателя электрона (1897) и нобелевского лауреата 1906 г. «за исследования прохождения электричества через газы».

Однако мы видим волны на поверхности реки, а вода состоит из частиц. Чтобы убедиться, что дифракция электронов не является коллективным эффектом, советский ученый В. А. Фабрикант осуществил красивый опыт (1949). У него время пролета от электронной пушки до мишени ( $7 \cdot 10^{-9}$  с) было много меньше времени между последовательными вылетами электронов ( $2,4 \cdot 10^{-4}$  с), т. е. электроны друг на друга никак не действовали. Тем не менее на экране возникала дифракционная картина, получаемая при падении на экран множества *одиночных* электронов. Это доказывает, что волновые свойства присущи не коллективу



**Рис. 35.5.** Фотография дифракционной картины рассеяния электронов, полученная Дж. П. Томсоном (из книги: В. Гейзенберг Физические принципы квантовой теории. М.: Гостехиздат, 1932)

электронов, а каждому из них в отдельности. И в то же время электрон — частица с зарядом и массой, оставляющая точечный след на фотопластинке.

Вопрос, долго волновавший физиков: что такое микрочастица — волна или корпускула? Здесь налицо некорректная постановка вопроса, подразумевающая альтернативу. Надо «или» заменить на «и»: электрон *и* волна, *и* частица. Если задать этот вопрос природе с помощью прибора, то получим ответ, соответствующий прибору: для волнового прибора (например, дифракционной решетки) ответ будет «волна», для корпускулярного (скажем, счетчика) — «частица». Неоднозначность ответа отражает двойственность природы частиц или, точнее, узость нашего альтернативного мышления, допускающего лишь одну из этих возможностей. Это новые для нас объекты, квантовые русалки и кентавры микромира, которых невозможно разделить на людей и животных. В этом смысле микрообъект не является ни волной, ни частицей и ни их симбиозом. Это — новое качество, которое мы количественно отражаем в формуле де Бройля (35.22) связывающей волновые и корпускулярные проявления свойств единого квантового объекта.

В экспериментах с ускоренными частицами их длина волны де Бройля играет ту же роль, что и длина волны видимого света, определяющая разрешающую силу микроскопа. Ускорители — своеобразные микроскопы микромира, и с их помощью исследуются законы природы на малых расстояниях. На Большом адронном коллайдере в Женеве энергия ускоряемых протонов достигает 7 ТэВ ( $7 \cdot 10^{12}$  эВ), что соответствует длине волны де Бройля  $\lambda \approx 2 \cdot 10^{-19}$  м (т.е. в миллиард раз меньше размеров атомов).

При решении задач, связанных с длиной волны де Бройля, надо с осторожностью применять общую формулу  $\lambda = h/p$ . Необходимо ясно понимать, какое выражение для импульса следует использовать. Например, в задаче может быть задана скорость частицы  $v$ . Если эта скорость много меньше скорости света  $c = 2,998 \cdot 10^8$  м/с, то можно применить классическое соотношение  $p = mv$ . (На практике считают, что  $v \ll c$ , если отношение  $v/c \leq 0,3$ , т.е.  $v^2/c^2 \leq 0,1$ .) Если же это не так, то следует применить *релятивистское* соотношение между импульсом частицы и ее скоростью:

$$p = \frac{mv}{\sqrt{1 - v^2/c^2}}.$$

В задаче может быть задана не скорость частицы, а ее кинетическая энергия  $E_k$ . Критерием применимости классических формул является малость кинетической энергии по сравнению с энергией покоя частицы  $E_0 = mc^2$ . Если условие  $E_k \ll mc^2$  выполняется (что, как нетрудно видеть, эквивалентно условию  $v \ll c$ ), то для нахождения импульса можно воспользоваться формулами классической механики

$$E_k = \frac{mv^2}{2} = \frac{p^2}{2m}, \quad \text{откуда} \quad p = \sqrt{2mE_k}.$$

Если же кинетическая энергия частицы не мала по сравнению с энергией покоя (тем более превышает ее), то следует использовать релятивистские формулы. В теории относительности выводится общее соотношение между полной энергией частицы  $E$  и ее импульсом, которое уже встречалось выше (см. (34.2)):

$$E = \sqrt{m^2c^4 + p^2c^2}, \quad \text{откуда} \quad p = \frac{1}{c} \sqrt{(E - mc^2)(E + mc^2)}. \quad (35.23)$$

Кинетическая энергия  $E_k$  равна разности между полной энергией и энергией покоя:  $E_k = E - mc^2$ , так что для импульса получаем

$$p = \frac{1}{c} \sqrt{E_k(E_k + 2mc^2)}.$$

Наконец, удобно переписать это выражение так, чтобы под знаком квадратного корня стояли безразмерные величины:

$$p = mc \sqrt{\frac{E_k}{mc^2} \left( \frac{E_k}{mc^2} + 2 \right)}. \quad (35.24)$$

Нетрудно убедиться, что при  $E_k \ll mc^2$  формула (35.24) действительно переходит в классическое выражение  $p = \sqrt{2mE_k}$ . Традиционные единицы энергии (Дж) оказываются неудобными в микромире. Поэтому физики предпочитают пользоваться несистемными единицами (мы ими уже пользовались в этой главе) — электронвольт (эВ) и его производными (1кэВ =  $10^3$ эВ, 1МэВ =  $10^6$ эВ, 1ГэВ =  $10^9$ эВ и т. п.).

Электронвольт — это энергия, приобретаемая электроном при прохождении разности потенциалов 1 В:  $E = eU$ . Заряд электрона равен  $e = 1,602 \cdot 10^{-19}$  Кл, так что 1 эВ =  $1,602 \cdot 10^{-19}$  Дж. Поскольку у нас появилась энергия покоя, приведем ее численные значения для электрона и протона вместе с массами этих частиц. Заодно укажем численные значения постоянной Планка в соответствующих единицах — эВ · с:

$$\begin{aligned} h &= 4,136 \cdot 10^{-15} \text{ эВ} \cdot \text{с}, & \hbar &= 6,582 \cdot 10^{-16} \text{ эВ} \cdot \text{с}, \\ m_e &= 9,109 \cdot 10^{-31} \text{ кг}, & m_e c^2 &= 8,187 \cdot 10^{-14} \text{ Дж} = 0,511 \text{ МэВ}, \\ m_p &= 1,673 \cdot 10^{-27} \text{ кг}, & m_p c^2 &= 1,503 \cdot 10^{-10} \text{ Дж} = 938,3 \text{ МэВ}. \end{aligned}$$

Наконец, встречаются обратные задачи, когда задана длина волны де Бройля  $\lambda$  и требуется найти скорость или энергию частицы. Возникает вопрос: как по

величине  $\lambda$  сразу определить, какими формулами следует пользоваться? Представим себе, что мы применяем для нахождения скорости классическую формулу  $p = mv$  и, следовательно, получаем  $v = h/(m\lambda)$ . Критерием справедливости этого результата является малость скорости  $v$  по сравнению со скоростью света  $c$ :

$$\frac{h}{m\lambda} \ll c, \text{ откуда } \lambda \gg \frac{h}{mc}.$$

Как уже говорилось, комбинации  $\lambda_C = h/(mc)$  и  $\lambda_C = \hbar/(mc)$  называются комптоновской длиной волны частицы. По своему физическому смыслу эта величина задает характерные расстояния, на которых уже неприменима нерелятивистская квантовая механика. Приведем численные значения этого важного параметра для электрона и протона:

$$\begin{aligned} \lambda_C^{(e)} &= \frac{h}{m_e c} = 2,426 \cdot 10^{-12} \text{ м}, & \lambda_C^{(e)} &= \frac{\hbar}{m_e c} = 3,861 \cdot 10^{-13} \text{ м}, \\ \lambda_C^{(p)} &= \frac{h}{m_p c} = 1,321 \cdot 10^{-15} \text{ м}, & \lambda_C^{(p)} &= \frac{\hbar}{m_p c} = 2,103 \cdot 10^{-16} \text{ м}. \end{aligned} \quad (35.25)$$

Видно, что чем легче частица, тем больше ее комптоновская длина волны, т. е. для тяжелых частиц релятивистские эффекты проявятся на меньших расстояниях, чем для легких. Для электрона комптоновская длина волны много меньше характерной длины, определяющей размеры атома. Это значит, что к электронам в атоме применима нерелятивистская квантовая механика. Ядра, состоящие из протонов и нейтронов, имеют размеры порядка  $10^{-14}$  м, что много больше комптоновской длины волны протона. Следовательно, к ним тоже применима эта теория. Условие нерелятивизма можно записать также в виде  $\lambda \gg \lambda_C$ , где  $\lambda$  — длина волны де Бройля.

**Задача 35.42.** Определить скорость частицы, у которой длина волны де Бройля в 10 раз меньше комптоновской длины волны  $\lambda_C$ .

**Решение.** По условию

$$\frac{h}{mc} = 10 \frac{h}{p},$$

откуда находим для импульса частицы:

$$p = 10mc.$$

Так как длина волны де Бройля меньше комптоновской длины волны частицы, то мы не можем воспользоваться нерелятивистским соотношением  $p = mv$ , которое приведет к несуразному ответу  $v = 10c$  (как известно, ничто не может двигаться быстрее света). Здесь необходимо применить релятивистскую формулу связи импульса со скоростью:

$$p = \frac{mv}{\sqrt{1 - v^2/c^2}}.$$

Отсюда вытекает уравнение

$$\frac{v/c}{\sqrt{1 - v^2/c^2}} = 10,$$

решение которого дает

$$v = \frac{10}{\sqrt{101}}c = 0,995c,$$

т. е. скорость частицы всего на 0,5% меньше скорости света. ■

### 35.3 Соотношения неопределенностей

Наличие у частицы волновых свойств накладывает определенные ограничения на возможность корпускулярного описания ее поведения. Для классической частицы всегда можно указать ее точное положение и импульс. Для квантового объекта имеем иную ситуацию.

Представим цуг волн пространственной протяженностью  $\Delta l$  — образ локализованного электрона, положение которого известно с точностью до  $\Delta l$ . Длину волны электрона можно определить, подсчитав число периодов на отрезке:  $N = \Delta l/\lambda$ . Какова точность определения  $\lambda$ ? Ясно, что для слегка различающейся длины волны мы получим то же самое значение  $N$ . Неопределенность  $\Delta\lambda$  в длине волны ведет к неопределенности  $\Delta N = \Delta l(\Delta\lambda/\lambda^2)$  в числе узлов, причем измерению доступны лишь  $\Delta N \geq 1$ . Так как  $p = h/\lambda$  и  $\Delta p = h\Delta\lambda/\lambda^2$ , отсюда немедленно следует соотношение

$$\Delta l \Delta p \geq h.$$

Точности ради надо заметить, что, во-первых, величина  $p$  в данном случае означает проекцию импульса на направление  $l$  и что, во-вторых, приведенное рассуждение имеет скорее качественный, нежели количественный, характер, поскольку мы не дали строгой математической формулировки, что понимается под неопределенностью измерения. Если это сделать математически строго, то придем к трем неравенствам (по числу координатных осей):

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}, \quad \Delta y \Delta p_y \geq \frac{\hbar}{2}, \quad \Delta z \Delta p_z \geq \frac{\hbar}{2}. \quad (35.26)$$

Это и есть знаменитые *соотношения неопределенностей Гейзенберга для координат–импульсов* (1927), уступающие по популярности среди широкой публики лишь формуле  $E = mc^2$ .

Представим теперь, что мы стоим на месте и мимо проходит электронная волна. Наблюдая за ней в течение времени  $\Delta t$ , хотим найти ее частоту  $\nu$ . Насчитав  $N = \nu\Delta t$  колебаний, определим частоту с точностью до  $\Delta\nu = \Delta N/\Delta t \geq 1/\Delta t$ , откуда имеем  $\Delta\nu\Delta t \geq 1$  или (с учетом формулы  $E = h\nu$ ) получаем:

$$\Delta E \Delta t \geq h.$$

*Соотношение неопределенностей Гейзенберга для энергии–времени* по причинам, упомянутым выше, чаще используется в виде

$$\Delta E \Delta t \geq \frac{\hbar}{2}. \quad (35.27)$$

Поговорим о физическом смысле этих соотношений. Здесь снова проявляется дополнительность макроскопических приборов. Но приборы совсем не виноваты: ограничения носят принципиальный, а не технический характер. Они существуют, даже если мы не проводим никаких экспериментов по измерению положения

и импульса частицы. Сам микрообъект не может быть в состоянии, для которого определенные значения одновременно имеют какая-то из его координат и проекция импульса на ту же ось.

Смысл второго соотношения: если микрообъект живет конечное время, то его энергия не имеет точного значения, она как-бы размыта. **Естественная ширина спектральных линий** — прямое следствие формул Гейзенберга. В модели атома Бора предполагается, что время жизни стационарных состояний бесконечно — в этом и состоит физический смысл понятия стационарного состояния. Соответственно их энергия определена точно. На самом деле в возбужденном состоянии атом в среднем живет конечное время, после чего возвращается в более низкое состояние. Следовательно, энергия возбужденных уровней точно не определена. Значит, некоторый разброс имеет и энергия излучаемого при переходе фотона, что приводит к уширению спектральной линии<sup>1</sup>.

Имеется и другое проявление соотношения неопределенностей энергия–время. Если неопределенность в энергии электрона превышает расщепление между энергиями соседних состояний  $\Delta E \geq (E_1 - E_2)$ , то нельзя точно сказать, на каком уровне находится электрон. Иными словами, на короткое время порядка  $\Delta t \approx \hbar/(E_1 - E_2)$  электрон может перескочить с уровня 1 на уровень 2, не излучая фотона, и затем вернуться назад. Это — так называемый **виртуальный процесс**, который в принципе не наблюдается и, следовательно, не нарушает закона сохранения энергии.

Соотношения неопределенностей имеют место и для других пар так называемых *сопряженных динамических переменных*. Так, при вращении частицы вокруг некоторой оси по орбите радиусом  $r$  неопределенность угловой координаты частицы  $\Delta\varphi$  влечет за собой неопределенность ее положения на орбите  $\Delta x = r\Delta\varphi$ . Из соотношений (35.26) следует, что неопределенность импульса частицы удовлетворяет неравенству  $\Delta p \geq \hbar/(r\Delta\varphi)$ . Учитывая связь момента количества движения  $L$  с импульсом ( $L = pr$ ), получаем  $\Delta L = r\Delta p$ , откуда следует соотношение неопределенностей

$$\Delta L \Delta\varphi \geq \frac{\hbar}{2}. \quad (35.28)$$

## Некоторые следствия соотношений неопределенностей

### Отсутствие траекторий у микрочастиц

Для нерелятивистской частицы  $p = mv$  и

$$\Delta t \Delta v \geq \frac{\hbar}{2m}.$$

Для массивных объектов правая часть исчезающе мала, что позволяет одновременно измерить скорость и положение объекта (область справедливости классической механики). В атоме же Бора импульс электрона  $p = \hbar/a_B$ , и неопределенность положения оказывается порядка размеров орбиты!

<sup>1</sup>Заметим, что соотношение  $\Delta\nu\Delta t \geq 1$  между разбросом частот в сигнале и его длительностью давно известно в радиотехнике.



**Невозможность достижения минимума потенциальной энергии**

Рассмотрим осциллятор с энергией

$$E = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2}.$$

Основное состояние в классической механике:  $p = 0, x = 0$ . Поэтому величины неопределенностей импульса и координаты — порядка самих импульса и координаты, т. е.:

$$p \approx \frac{\hbar}{2x} \rightarrow E = \frac{\hbar^2}{8mx^2} + \frac{m\omega^2 x^2}{2}.$$

Минимум энергии достигается в точке  $x = \sqrt{\hbar/(2m\omega)}$  и равен

$$E = \frac{\hbar\omega}{2}.$$

Действуя таким образом, мы получаем лишь оценку энергии основного состояния по порядку величины, но не ее точное значение (хотя ответ, полученный нами, в данном случае действительно точен). Мы только что обнаружили **нулевые колебания**: квантовый осциллятор, в отличие от классического, не может оставаться в покое в точке минимума потенциальной энергии — это противоречит соотношению неопределенностей Гейзенберга. Мы увидим в дальнейшем, что этот вывод имеет важные следствия в самых разных областях физики. Точные расчеты показывают, что формулу Планка для уровней энергии осциллятора следует писать в виде

$$E = \hbar\omega \left( n + \frac{1}{2} \right), \quad n = 0, 1, 2, 3, \dots \quad (35.29)$$

При решении задач на применение соотношения неопределенностей следует иметь в виду, что в основном состоянии в классической физике электрон покоится в точке, соответствующей минимуму потенциальной энергии. Соотношения неопределенностей не позволяют ему это делать в квантовой теории, так что электрон должен иметь некоторый разброс импульсов. Поэтому неопределенность импульса (его отклонение от классического значения 0) и сам импульс по порядку величины совпадают ( $p \sim \Delta p$ ). Неопределенность координаты электрона, замкнутого в объеме  $V$ , равна по порядку величины линейному размеру этого объема ( $\Delta x \sim \sqrt[3]{V}$ ).

**Задача 35.43.** Электрон с кинетической энергией  $E_k = 15$  эВ находится на металлической пылинке диаметром  $d = 1$  мкм. Оценить относительную неопределенность  $\delta v$ , с которой может быть определена скорость электрона.

**Решение.** По условию  $\Delta x \approx d/2$  и  $\Delta p \approx 2\hbar/d$ . Энергия электрона много меньше его энергии покоя, так что можно применять нерелятивистские соотношения  $E_k = m_e v^2/2$ ,  $p = \sqrt{2E_k/m_e}$ ,  $\Delta p = m_e \Delta v$ , откуда  $\Delta v \approx 2\hbar/(m_e d)$ . Для относительной неопределенности скорости электрона получаем

$$\delta v = \frac{\Delta v}{v} \approx \frac{2\hbar}{d\sqrt{2m_e E_k}} = \frac{\hbar}{m_e c d} \sqrt{\frac{2m_e c^2}{E_k}} = \frac{\lambda_C^{(e)}}{d} \sqrt{\frac{2m_e c^2}{E_k}}.$$

В этом выражении выделены безразмерные множители — отношение комптоновской длины волны электрона к диаметру пылинки и отношение энергии покоя электрона

к его кинетической энергии. Подставляем численные значения:

$$\delta v = \frac{3,861 \cdot 10^{-13}}{1 \cdot 10^{-6}} \sqrt{\frac{2 \times 0,511 \cdot 10^6}{15}} \approx 10^{-4}. \blacksquare$$

**Задача 35.44.** Среднее время жизни атома в возбужденном состоянии  $\tau = 10^{-8}$  с. При переходе в основное состояние излучается фотон с длиной волны  $\lambda = 0,5$  мкм. Определить энергию кванта, ширину и относительную ширину спектральной линии.

**Решение.** Энергия кванта равна:

$$E = h\nu = \frac{hc}{\lambda} = \frac{4,136 \cdot 10^{-15} \times 3 \cdot 10^8}{0,5 \cdot 10^{-6}} \text{ эВ} = 2,48 \text{ эВ}.$$

Неопределенность в энергии определяется из соотношения (35.27), где  $\Delta t \sim \tau$ :

$$\Delta E = \frac{\hbar}{2\tau} = \frac{6,582 \cdot 10^{-16}}{2 \cdot 10^{-8}} \text{ эВ} = 3,3 \cdot 10^{-8} \text{ эВ}.$$

Отсюда

$$\frac{\Delta \lambda}{\lambda} = \frac{\Delta E}{E} = 1,3 \cdot 10^{-8},$$

$$\Delta \lambda = 1,3 \cdot 10^{-8} \times 0,5 \text{ мкм} = 0,7 \cdot 10^{-8} \text{ мкм}.$$

Это и есть естественная ширина спектральной линии, о которой упоминалось выше.  $\blacksquare$

Выдающийся физик Георгий Гамов (один из основоположников теории «горячей Вселенной», помимо всего прочего) написал цикл рассказов о том, как некий мистер Томпкинс попадает в миры, где мала скорость света или велика постоянная Планка<sup>2</sup>. Таким образом в увлекательной форме Гамов популяризировал теорию относительности и квантовую механику, рассказывая просто о необычайных и сложных явлениях, с которыми пришлось столкнуться любопытному мистеру Томпкинсу. В одной из глав тот попадает в квантовые джунгли и отправляется на охоту верхом на слоне. На них нападает стая тигров, размазанных в пространстве (рис. 35.6, слева). Но это не множество тигров, как было подумал мистер Томпкинс, — это один и тот же тигр, местоположение которого невозможно точно локализовать из-за соотношения неопределенностей. Далее охотники видят, как из бамбуковых зарослей выскакивает стадо газелей, похожих друг на друга, как две капли воды (рис. 35.6, справа). На самом деле это одна газель, но вследствие своих волновых свойств она «размножилась» после прохождения дифракционной решетки, роль которой сыграли стволы бамбука. К нашему удовольствию, из-за малости постоянной Планка подобные явления наблюдаются лишь в микромире.

**Задача 35.45.** На лестнице стоит человек и с высоты  $H$  роняет вниз шарики массой  $m$ , стараясь попасть в центр помещенной под ним мишени. Предполагается, что он целится идеально точно, размерами шарика и сопротивлением воздуха также пренебрегаем. Используя принцип неопределенностей Гейзенберга, оценить средний разброс шариков около центра мишени.

<sup>2</sup>Г. Гамов. Приключения мистера Томпкинса. Библиотечка «Квант». М.: Бюро Квантум, 1993. На английском языке книга вышла в 1940 г., более чем на полвека раньше, чем в нашей стране. Это связано с тем, что Гамов, ставший в возрасте 28 лет самым молодым членом-корреспондентом АН СССР, в 1933 г. не вернулся в страну из зарубежной поездки и в конечном итоге осел в США.



Рис. 35.6. Иллюстрации из научно-популярной книги Г. Гамова «Приключения мистера Томпкинса»

**Решение.** Направим ось  $Oy$  вертикально вниз, ось  $Ox$  горизонтально, а начало отсчета системы координат поместим в точке бросания. При свободном падении тела в поле силы тяжести  $y = gt^2/2$ , откуда имеем для времени падения  $t_{\text{пад}} = \sqrt{2H/g}$ .

Начальное положение шарика по горизонтали можно определить лишь с точностью до  $\Delta x$ . Это приводит к погрешности в горизонтальной составляющей импульса, минимальное значение которой определяется соотношением неопределенностей:  $\Delta p_x = \hbar/(2\Delta x)$ . Соответствующая горизонтальная компонента скорости равна  $\Delta v_x = \Delta p_x/m$ . Поэтому со временем горизонтальное смещение шарика меняется по закону

$$x(t) = \Delta x + \frac{\Delta p_x}{m}t = \Delta x + \frac{\hbar}{2m\Delta x}t.$$

Подставляя сюда время падения, находим смещение упавшего шарика относительно центра мишени:

$$x(t_{\text{пад}}) = \Delta x + \frac{\hbar}{2m} \sqrt{\frac{2H}{g}} \frac{1}{\Delta x}. \quad (35.30)$$

Минимальное расстояние  $s$  упавшего шарика от отвеса, проведенного к центру мишени, определяется дифференцированием (см. (35.30)). Получаем:

$$\frac{dx(t_{\text{пад}})}{d(\Delta x)} = 1 - \frac{\hbar}{2m} \sqrt{\frac{2H}{g}} \frac{1}{(\Delta x)^2} = 0, \quad (35.31)$$

откуда

$$(\Delta x)^2 = \frac{\hbar}{m} \sqrt{\frac{H}{2g}}. \quad (35.32)$$

Подставляя (35.32) в (35.30), находим минимальное расстояние  $s$  упавших шариков от центра мишени:

$$s = 2 \left( \frac{\hbar^2 H}{2m^2 g} \right)^{1/4}. \quad (35.33)$$

Численная оценка для шариков массой  $m = 10$  мг, бросаемых с высоты  $H = 10$  м:

$$s = 2 \left( \frac{(1,055 \cdot 10^{-34})^2 \times 10}{2 \times (10^{-2})^2 \times 9,81} \right)^{1/4} = 0,87 \cdot 10^{-16} \text{ м.}$$

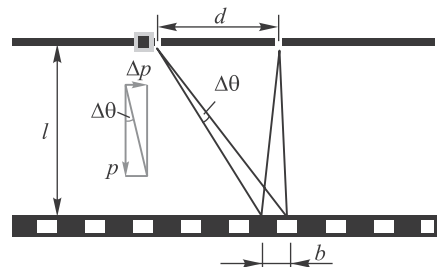
Такой разброс меньше радиуса протона. ■

## 35.4 Двухщелевой эксперимент

Опыт Юнга по интерференции света от двух когерентных источников (прорезей в защитном экране) обсуждался в разделе 29.4. Коль скоро электроны имеют волновые свойства, нас уже не удивит, что аналогичная интерференционная картина должна наблюдаться (и наблюдается) при падении на подобные «щели» электронов, протонов, нейтронов и даже очень крупных молекул. Все они обладают свойствами и волны, и частицы.

Но одиночный электрон оставляет на фотопластинке точку, а интерференционная картина возникает при падении множества электронов, пусть даже они подлетают к щелям поодиночке. Можно было бы представить себе, что одиночные электроны, будучи обычными частицами, все же проходят через какую-то одну из двух щелей, но не через обе сразу, как это происходит с обычной классической волной. Пусть нам неизвестно, по каким закономерностям электроны «выбирают» свои пути, но предположим, что эти пути существуют и выливаются в конечном итоге в интерференционную картину. Коль скоро это так, можно попытаться «подсмотреть», через какую именно щель пройдет тот или иной электрон, установив около нее какой-то пролетный детектор. Конструкция его нам неважна, сейчас мы занимаемся тем, что физики называют *мысленным экспериментом*. Сам термин (нем. *Gedankenexperiment*) ввел знаменитый немецкий ученый и философ Эрнст Мах. Согласно Википедии, «мысленный эксперимент в физике зачастую напоминает доказательство теоремы методом от противного в математике, когда некоторое положение физической модели или схемы сначала отвергается, а затем путем преобразования модели мы приходим к противоречию с тем или иным принципом, который считается, безусловно, истинным».

Итак, предположим, что у одной из щелей мы установили детектор, способный зафиксировать прохождение электрона (рис. 35.7). При измерении положения электрона детектор вносит неопределенность  $\Delta p$  в поперечную компоненту импульса электрона. Это приводит к повороту импульса на угол  $\Delta\theta = \Delta p/p$  и соответственно к сдвигу интерференционной картины на расстояние  $b = l\Delta\theta$ , где  $l$  — расстояние от щелей до



**Рис. 35.7.** Детектор, способный определить, через какую щель прошел электрон, разрушает интерференционную картину

■ максимумы    ■ детектор  
□ минимумы

экрана (фотопластинки). Комбинируя эти соотношения, получаем для сдвига интерференционного узора

$$b = l \frac{\Delta p}{p}. \quad (35.34)$$

Из оптики известно, что расстояние между максимумами и минимумами интерференционной картины равно  $s = \lambda l / (2d)$ , где  $d$  — расстояние между щелями. Интерференционная картина сохранится, если сдвиг полос  $b$ , вызванный наличием детектора, много меньше расстояния между полосами:  $b \ll s$ . Отсюда получаем

$$\Delta p \ll \frac{\lambda p}{2d} = \frac{h}{2d}. \quad (35.35)$$

Здесь мы использовали выражение (35.22) для длины волны де Бройля. Тогда из соотношения неопределенностей следует ограничение на точность определения детектором положения электрона:

$$\Delta x \geq \frac{\hbar}{2\Delta p} \gg \frac{d}{2\pi}. \quad (35.36)$$

Знак «много больше» в физике означает «больше на порядок», т. е. в десять (минимум в пять) раз. Таким образом, детектор, не меняющий интерференционную картину, способен определить положение электрона с точностью до  $\Delta x$  порядка расстояния  $d$  между щелями, т. е. он не способен различить, через какую именно щель прошел электрон.

Мы пришли к парадоксальному выводу: электрон проявляет волновые свойства, только если за ним не подглядывают. То же самое относится к фотонам или любым микрочастицам. Двухщелевые эксперименты проводятся не только мысленно, но и в реальности, причем в изощренном виде. Удивительные результаты получены в экспериментах с *квантовым «ластиком»*. Детекторы, позволяющие установить щель, через которую прошла частица, уничтожают интерференционную картину. Но если между щелями и экраном стоит устройство, стирающее эту информацию, то интерференционная картина восстанавливается, хотя, казалось бы, «ластик» действует уже после прохождения щелей частицами. Подобные эксперименты, проведенные уже в наши дни, демонстрируют правильность представления о корпускулярно-волновом дуализме, который Р. Фейнман назвал «единственной настоящей тайной квантовой механики».

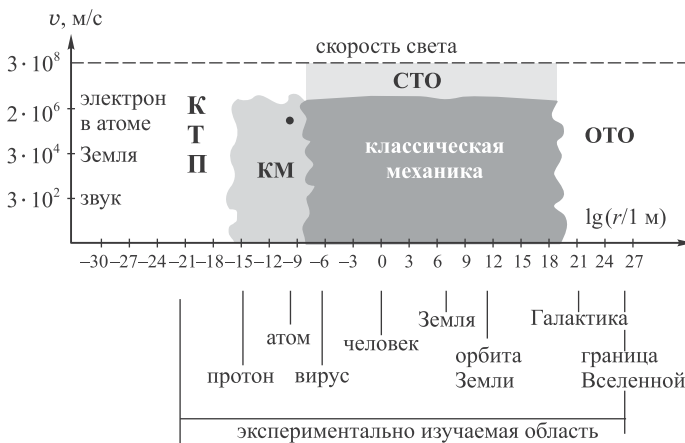
## 35.5 О границах применимости квантовой механики

Собственно, мы еще с квантовой механикой даже не познакомились, но лишь подошли к ее порогу. Тем не менее уже сейчас можно дать оценку области, в которой заведомо не будут справедливы выводимые нами соотношения. Если подходить к микромиру со стороны области применимости классической физики, то никаких проблем не возникает. В самом деле, квантовые соотношения вовсе не отменяют классических законов, но уточняют их. В макромире с большими значениями энергий и моментов количества движения квантовая дискретность просто будет незаметна, так что спокойно можно применять и квантовые, и классические законы. Разница между получаемыми с их помощью количественными результатами будет ничтожно мала, но классические законы проще. В этом смысле мы говорим об области, где справедливы законы классической физики.

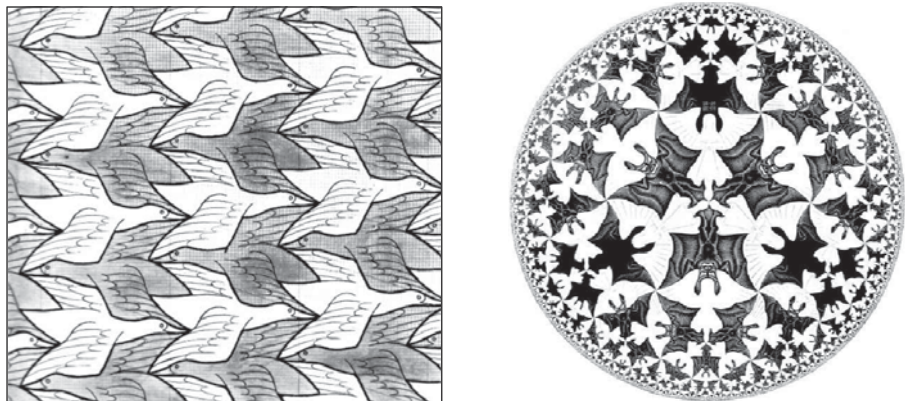
Иное дело, когда мы станем продвигаться в глубь микромира. В принципе ясно, что где-то откажут и законы квантовой механики, и ученым придется придумать какую-то новую теорию. В этом курсе мы будем заниматься нерелятивистской квантовой механикой, изучающей микрообъекты, движущиеся со скоростями, гораздо меньшими скорости света  $c$ . Мы встретимся и с релятивистскими объектами (фотонами и т. д.), но будем оговаривать такие случаи особо.

Определим теперь область справедливости нерелятивистской квантовой механики. Представим себе микрочастицу с массой  $m$ . Квантовая неопределенность  $\Delta E$  ее энергии должна быть много меньше энергии покоя  $mc^2$  (иначе квантовые эффекты будут релятивистскими, что потребует обобщения теории). Из соотношения неопределенностей (35.27) и условия применимости нерелятивистских уравнений  $\Delta E \ll mc^2$  следует ограничение на интервалы времени:  $\Delta t \gg \hbar/(mc^2)$ . Характерное время  $t_c = \hbar/(mc^2)$  является, таким образом, водоразделом между релятивистской и нерелятивистской квантовыми теориями. За это время частица может пройти расстояние не больше  $ct_c = \hbar/(mc)$ . Величина  $\hbar/(mc)$  — характерное расстояние, на котором проходит граница между нерелятивистской и релятивистской теориями. Для электрона оно равно  $3,86 \cdot 10^{-13}$  м, а для протона —  $2,1 \cdot 10^{-16}$  м. Соответствующие промежутки времени равны  $t_c = 1,3 \cdot 10^{-21}$  с для электрона и  $t_c = 7 \cdot 10^{-25}$  с для протона.

На рисунке 35.8 по оси  $x$  в логарифмическом масштабе отложены расстояния, а по оси  $y$  — скорости. Отмечены области применимости основных физических теорий (ОТО — общая теория относительности, СТО — специальная теория относительности, КТП — квантовая теория поля, она же — релятивистская квантовая теория, она же — теория элементарных частиц). Штриховкой показаны области



**Рис. 35.8.** Диаграмма расстояние—скорость для наглядной иллюстрации областей применимости физических теорий: классической механики, специальной теории относительности (СТО), квантовой механики (КМ), общей теории относительности (ОТО) и квантовой теории поля (КТП); по оси ординат отложена скорость частицы, по оси абсцисс — расстояние в логарифмической шкале; точка показывает электрон в атоме — в сфере ответственности квантовой механики



**Рис. 35.9.** Картины Маурица Эшера «Две птицы» и «Ангелы и черти» как художественный образ корпускулярно-волнового дуализма

применимости теорий, изучавшихся нами. Правая граница заштрихованной области КМ не означает, что квантовую механику нельзя применять там, где мы привыкли пользоваться классической. Просто это делать нецелесообразно.

Понятно, что четких границ между различными теориями нет, они размыты, так как теории не отрицают, а развивают и дополняют одна другую. Так, квантовая теория поля (КТП), занимающаяся свойствами элементарных частиц, родилась при соединении квантовой механики (КМ) и специальной теории относительности (СТО), а классическая теория гравитации — она же общая теория относительности ОТО — появилась при распространении СТО и классической механики на область больших расстояний.

Обратим внимание на то, что рисунок ограничен с обеих сторон. Большой взрыв, в котором родилась наша Вселенная, произошел 13,7 млрд лет тому назад. За это время свет преодолел путь порядка  $10^{26}$  м, что и определяет максимальные возможные расстояния, доступные для наблюдения. На расстояниях же порядка  $10^{-35}$  м необходимо учитывать силы гравитации, квантовая теория которых пока не создана. Только очень смелые теоретики решаются серьезно обсуждать, что происходит на столь малых расстояниях. Поэтому данная область на рисунке не показана.

Подведем итоги этой главы. Микрочастицы проявляют **корпускулярно-волновой дуализм**: в одних экспериментах они ведут себя как частицы, в других — как волны.

Художественное выражение идеи дуализма можно найти в картинах знаменитого голландского художника Маурица Эшера (1898–1972). На рисунке 35.9 представлены две его картины: слева — «Две птицы», справа — «Ангелы и черти». Что вы видите на левой картине — белых птиц на темном фоне (в оригинале синем) или темных птиц на белом фоне? Можно увидеть и то, и другое, а что на самом деле? Вопрос «а какие птицы на самом деле?» смысла не имеет, ответ «они одновременно и те, и другие» похож на ответы квантовой механики. Аналогичные иллюзии созданы художником и на второй картине.

Но следует отметить, что эти иллюзии являются коллективным эффектом, создаваемым всем множеством нарисованных птиц. Если мысленно вырезать из картины лишь одну птицу, она определенно будет либо синей, либо белой. В квантовой же механике любой микрообъект является одновременно и частицей, и волной, но квантовые свойства невозможно изобразить в макроскопических масштабах привычного нам мира.

Однако на данном этапе вы пришли к противоречивой картине микромира. Модель атома Бора предсказала правильный спектр атома водорода. Но при расчете энергетических уровней использовались законы движения по классической траектории, а позже выяснилось, что никаких траекторий у квантовых объектов быть не может.

Если никаких других орбит, кроме стационарных, у электрона не существует, то он должен перескакивать с одной орбиты на другую мгновенно, и никакого пути перескока у него быть не должно — электрон просто не может находиться между стационарными орбитами. Более того, он не может даже находиться в какой-то определенной точке стационарной орбиты, так что совершенно непонятно, откуда и куда он перескакивает.

Эти мгновенные квантовые скачки сильно смущали в свое время даже самых выдающихся физиков. Даже Эрвин Шрёдингер, один из создателей квантовой механики, однажды в отчаянии воскликнул: «Если мы собираемся сохранить эти проклятые квантовые скачки, то я вообще сожалею, что имел дело с квантовой теорией!» На что Нильс Бор ответил ему: «Зато остальные весьма признательны Вам за это, ведь благодаря Вам был сделан решающий шаг в развитии квантовой теории».

Классический расчет радиуса орбиты в атоме противоречит волновой природе электрона, но именно подобные расчеты привели к связи его импульса и длины волны де Бройля.

Наконец, если электрон в каких-то обстоятельствах ведет себя как волна, то совершенно непонятно, что же колеблется и распространяется в пространстве, т. е. какова физическая природа этих волн материи.

## Контрольные вопросы

1. Что такое спектральные серии?
2. Сформулируйте постулаты Бора.
3. Какие значения могут принимать радиусы электронных орбит в атоме водорода по теории Бора?
4. Какие значения может принимать скорость электрона, находящегося в поле ядра с зарядом  $+Ze$ ? На основании полученного результата покажите, что в природе не могут существовать ядра с  $Z \geq 137$ .
5. Какие значения может принимать энергия электрона в атоме водорода по теории Бора?
6. Как теория Бора объясняет спектральные серии?
7. В чем сущность гипотезы де Бройля?



8. Какова длина волны де Бройля у электрона с кинетической энергией, равной его энергии покоя?
9. Сформулируйте соотношения неопределенностей Гейзенберга.
10. Что такое нулевые колебания?
11. По какой причине спектральные линии имеют конечную ширину?
12. Вообразите мир, где постоянная Планка имеет значение  $\hbar = 2000$  Дж · с. Живя в этом мире, вы назначаете свидание, которое может состояться, если вы вдвоем окажетесь в одном месте — скажем, на расстоянии порядка  $\Delta x \sim 1$  м друг от друга. Будет ли у вас возможность спокойно пообщаться на свидании?

## Глава 36

# Уравнение Шрёдингера

Вот мы и подошли к настоящей квантовой механике. Все, что было до сих пор, — это интуитивные полуклассические представления, позволившие инкорпорировать в классическую физику идеи физики квантовой. Но этот уровень знаний недостаточен для расчетов, количественных предсказаний многих явлений. Требуется стройная система, теория движения (или распространения) микрочастиц с дуальными (волна—корпускула) свойствами. В этой главе мы сформулируем несколько правил, которые позволят осмыслить переход от классической механики к квантовой.

### 36.1 Волна вероятности

Предыдущая глава закончилась констатацией, что мы пока не установили, что именно колеблется при движении электрона. В истории физики такое уже случалось. Когда-то при выводе уравнений электродинамики Максвелл тоже не знал, что представляют собой описываемые ими колебания и волны, но уравнения оказались верны. Поэтому отложим пока вопрос о физической природе волн де Бройля и просто введем некую «электронную» волну, т. е. функцию  $\Psi(t, \vec{r})$ . О ней мы знаем пока только одно: волновая функция должна описывать результаты опытов, доказывающих волновые свойства электронов (дифракцию и т. п.).

Представим мысленно эксперимент В. А. Фабриканта, в котором электроны поочередно направлялись на кристалл, играющий роль дифракционной решетки. За кристаллом помещена фотопластинка, на которой в конце концов возникают типичные дифракционные кольца. Из классической физики известно, какие математические средства описывают такую картину: обычное сложение интерферирующих волн, интенсивности которых пропорциональны  $|\Psi(t, \vec{r})|^2$ .

Правда, в отличие от обычной волны электрон не делится на части: при прохождении электронов через кристалл каждый из них попадает в какую-то одну точку на фотопластинке, вызывая почернение именно здесь, и нигде больше. В этом проявляются свойства электрона как частицы. Несмотря на одинаковые начальные условия, электроны, как показал опыт, попадают в разные точки. О данном конкретном электроне заранее неизвестно, в какую именно точку на пластинке он попадет. В этом проявляются его волновые свойства. Дифракционная

картина возникает, когда через кристалл пройдет достаточно много таких электронов. Интенсивность почернения пластинки в данной точке пропорциональна числу попавших туда частиц, т. е. вероятности попадания.

В классической же физике почернение пластинки определяется интенсивностью волны, т. е. квадратом модуля волновой функции  $|\Psi(t, \vec{r})|^2$ . Выходит, что он пропорционален вероятности обнаружить электрон в точке  $\vec{r}$  в момент времени  $t$ . Волна де Бройля — это волна вероятности! Отдельный акт взаимодействия электрона с кристаллом остается отдельным актом (электрон — частица), но результат его можно предсказать только вероятностно, статистически (электрон — волна). В этом смысл корпускулярно-волнового дуализма. Квантовая механика создана в 1925–27 гг. немецким физиком В. Гейзенбергом (Нобелевская премия в 1932 г. «за создание квантовой механики, применение которой привело, помимо прочего, к открытию аллотропических форм водорода»<sup>1</sup>) и австрийским физиком Э. Шрёдингером (Нобелевская премия в 1933 г. «за открытие новых продуктивных форм атомной теории»). Вероятностная интерпретация волновой функции была дана чуть позже в работах М. Борна (Нобелевская премия в 1954 г. «за фундаментальные исследования по квантовой механике, особенно за статистическую интерпретацию волновой функции») и школы Н. Бора.

Итак, вероятность найти электрон вблизи точки  $\vec{r}$  должна быть пропорциональна  $|\Psi(t, \vec{r})|^2$ . Но вероятность обнаружить электрон точно в данном месте исчезающе мала; имеет смысл говорить лишь о его попадании в малый объем  $dV$ , окружающий эту точку. Ясно, что вероятность  $dW$  обнаружить там электрон пропорциональна величине объема. Поэтому для вероятности имеем

$$dW = |\Psi(t, \vec{r})|^2 dV. \quad (36.1)$$

Иными словами,  $|\Psi(t, \vec{r})|^2$  — *плотность вероятности* найти частицу в точке с координатой  $\vec{r}$ . Вероятность  $W_V$  найти частицу в каком-то конечном объеме  $V$  вычисляется с помощью сложения вероятностей, т. е. интегрированием

$$W_V = \int_V |\Psi(t, \vec{r})|^2 dV. \quad (36.2)$$

Интегрирование в (36.2) ведется по объему  $V$  (в случае одномерного движения — по отрезку). Полная вероятность найти частицу хоть где-нибудь в пространстве должна быть равна единице. Отсюда следует *условие нормировки волновой функции*: такой же интеграл по всему пространству равен единице, т. е.

$$\int |\Psi(t, \vec{r})|^2 dV = 1. \quad (36.3)$$

Замечание: выполнение этого условия возможно для тех задач, в которых классическая частица движется в ограниченной области пространства (финитное движение). Для других движений условие нормировки усложняется.

<sup>1</sup>Аллотропия — существование двух и более простых веществ одного и того же химического элемента, различных по строению и свойствам. В данном случае имеются в виду состояния орто- и пара-водорода.

Наблюдаемые физические величины должны описываться действительными числами и функциями. Соответственно мы представляли классические волны (звуковые, электромагнитные) в виде  $\xi = A \cos(\omega t - \vec{k} \cdot \vec{r} + \varphi_0)$ . Можно было бы воспользоваться математическим формализмом комплексных чисел, основываясь на формуле Эйлера:

$$e^{-i\varphi} = \cos \varphi - i \sin \varphi,$$

где  $i = \sqrt{-1}$  — мнимая единица. Тогда ту же волну можно было бы представить в виде действительной части выражений

$$\xi = \operatorname{Re}\Psi, \quad \Psi = C \exp[-i(\omega t - \vec{k} \cdot \vec{r})], \quad C = A e^{-i\varphi_0}. \quad (36.4)$$

Начальная фаза здесь включена в комплексную амплитуду  $C$ . В применении к классическим волнам оба формализма эквивалентны, так как в конечном итоге берется только действительная часть волны. В отличие от классических волн, волна вероятности комплексна. Наблюдаемые физические величины выражаются через квадрат модуля волновой функции, так что и в квантовой механике они будут описываться действительными числами. Но комплексность волновой функции имеет глубокую связь с законом сохранения электрического заряда, так что применение комплексных чисел и функций в квантовой механике — не прихоть, а необходимость. Поверхность постоянной фазы в волне  $\Psi$  (см. (36.4)) распространяется вдоль волнового вектора  $\vec{k}$ , фазовая скорость волны по-прежнему равна  $v = \omega/k$ , так что переход к комплексным волнам не меняет привычных нам соотношений.

## 36.2 Общее уравнение Шрёдингера

Волновая функция является главным объектом изучения в квантовой механике. Говоря о каком-то состоянии в классической физике, мы подразумевали, что в момент времени  $t = 0$  частица имела некие положение и скорость (импульс), а дальнейшая ее судьба была предопределена уравнениями движения Ньютона.

Состояние в квантовой механике имеет иной смысл: в момент времени  $t = 0$  задана волновая функция, изменение которой регулируется пока не известным нам уравнением (Шрёдингера). В этом смысле теперь понимается *причинность*: в классической механике — точные предсказания положений и скоростей, в квантовой механике — предсказания состояний (волновых функций). Уравнения новой физики (в данном случае уравнение Шрёдингера) никогда не выводятся логически из прежних принципов (иначе это будет не новая теория, а следствие старой). Но квантово-механическое уравнение должно иметь некие классические корни, поскольку классическая механика хороша в области своей применимости. Далее мы приведем не вывод, а наводящие соображения.

Свободной частице соответствует волна де Бройля, которую мы записываем в виде классической плоской волны:

$$\Psi(t, \vec{r}) = C e^{-i\omega t + i\vec{k} \cdot \vec{r}} = C e^{-\frac{i}{\hbar}(Et - \vec{p} \cdot \vec{r})}, \quad (36.5)$$

где модуль волнового вектора  $\vec{k}$  связан с длиной волны соотношением  $k = 2\pi/\lambda$ , а  $C$  — комплексная амплитуда волны. Мы использовали уже известную связь

энергии и импульса частицы с частотой и длиной волны де Бройля. Искомое уравнение для волновой функции не должно содержать  $E$  и  $\vec{p}$ , так как это характеристики конкретного состояния частицы. Попробуем найти операции над волновой функцией свободной частицы, позволяющие исключить параметры  $E$  и  $\vec{p}$ . Имеем для производных по времени и по пространственной координате  $x$ :

$$i\hbar \frac{\partial}{\partial t} \Psi(t, \vec{r}) = E \Psi(t, \vec{r}), \quad -i\hbar \frac{\partial}{\partial x} \Psi(t, \vec{r}) = p_x \Psi(t, \vec{r}). \quad (36.6)$$

Повторяя дифференцирование по координате  $x$ , получаем

$$-\hbar^2 \frac{\partial^2}{\partial x^2} \Psi(t, \vec{r}) = p_x^2 \Psi(t, \vec{r}). \quad (36.7)$$

Складывая (36.7) с аналогичными уравнениями для вторых производных по координатам  $y$  и  $z$ , приходим к соотношению

$$-\hbar^2 \Delta \Psi(t, \vec{r}) = \vec{p}^2 \Psi(t, \vec{r}), \quad (36.8)$$

где знаком  $\Delta$  обозначен оператор Лапласа:

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

В этом месте возникает различие между релятивистским и нерелятивистским случаями. Квантовая механика — *нерелятивистская теория*, в которой  $E = \vec{p}^2/2m$ . Это классическое соотношение позволяет связать дифференцирование по времени в (36.6) с дифференцированием по пространственным координатам в (36.8) и тем самым исключить из уравнения зависимость от энергии и импульса частицы:

$$i\hbar \frac{\partial}{\partial t} \Psi(t, \vec{r}) = -\frac{\hbar^2}{2m} \Delta \Psi(t, \vec{r}). \quad (36.9)$$

Это уравнение вполне бы нас устроило, но применимо оно пока только для свободной частицы. Легко понять, как должно выглядеть уравнение для системы с постоянным значением  $U_0$  потенциальной энергии. Полная энергия равна сумме

$$E = \frac{\vec{p}^2}{2m} + U_0,$$

так что получаем

$$i\hbar \frac{\partial}{\partial t} \Psi(t, \vec{r}) = -\frac{\hbar^2}{2m} \Delta \Psi(t, \vec{r}) + U_0 \Psi(t, \vec{r}). \quad (36.10)$$

Произвольную функцию можно аппроксимировать ступенчатой функцией, так что в окрестности каждой точки  $\vec{r}$  потенциальную энергию  $U(\vec{r})$  можно считать постоянной, и искомое уравнение почти с очевидностью следует из (36.10):

$$\boxed{i\hbar \frac{\partial}{\partial t} \Psi(t, \vec{r}) = -\frac{\hbar^2}{2m} \Delta \Psi(t, \vec{r}) + U(\vec{r}) \Psi(t, \vec{r}).} \quad (36.11)$$

Это и есть основное уравнение квантовой механики — знаменитое *общее уравнение Шрёдингера*. Подчеркнем еще раз, что вывести его строго невозможно, но можно угадать, исходя из наводящих соображений, что мы и сделали. Соответствие уравнения и его следствий физической реальности проверяется экспериментально.

Уравнение Шрёдингера для волновой функции — линейное дифференциальное уравнение в частных производных. Отсюда сразу же вытекают следствия:

- если  $\Psi(t, \vec{r})$  — решение уравнения (36.11), то  $A\Psi(t, \vec{r})$  также его решение при любой константе  $A$ , следовательно, подбором постоянной  $A$  можно добиться выполнения условия нормировки (36.3);
- если  $\Psi_1(t, \vec{r})$  и  $\Psi_2(t, \vec{r})$  — решения уравнения Шрёдингера, то линейная комбинация  $A_1\Psi_1(t, \vec{r}) + A_2\Psi_2(t, \vec{r})$  также его решение (иными словами, справедлив принцип суперпозиции, т. е. основа явления интерференции).

### 36.3 Операторы, симметрия и законы сохранения

Итак, состояние электрона описывается теперь волновой функцией  $\Psi(t, \vec{r})$ . Но куда подевались координаты, импульс и прочие величины, известные из классической теории? От классических представлений придется отказаться. Взамен у нас появились *дифференциальные операторы*, действующие на функцию  $\Psi(t, \vec{r})$ . Из уравнения Шрёдингера видно, что оно воспроизводит связь  $E = T + U$  полной энергии  $E$  с кинетической  $T$  и потенциальной  $U$ , но классические величины заменены на операторы, действующие на волновую функцию  $\Psi$ . Будем обозначать оператор тем же символом, что и классическую величину, снабжая его для отличия шляпкой. Тогда уравнение Шрёдингера (36.11) можно записать в операторной форме, в которой отчетливо видна его связь с энергетическими соотношениями классической физики:

$$\hat{E}\Psi = \hat{T}\Psi + \hat{U}\Psi,$$

где введены операторы

$$\hat{E} = i\hbar \frac{\partial}{\partial t}, \quad \hat{p} = -i\hbar \vec{\nabla}, \quad \hat{T} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m} \Delta. \quad (36.12)$$

Здесь  $\vec{\nabla}$  — оператор градиента с компонентами

$$\vec{\nabla} = \left\{ \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\},$$

квадрат которого дает оператор Лапласа  $\Delta$ . Оператор координаты  $\hat{r}$  сводится к простому умножению  $\Psi$  на вектор  $\vec{r}$ ; это справедливо для оператора любой функции  $\vec{r}$  (в частности, для потенциальной энергии).

Мы пришли к правилу квантования, т. е. к способу перехода от известной классической системы к соответствующей ей квантовой.

**Правило 1.** ♦ Классическим динамическим переменным  $(\vec{r}, \vec{p})$  в квантовой механике соответствуют оператор  $\hat{r}$  — умножение на вектор  $\vec{r}$  и оператор  $\hat{p} = -i\hbar\vec{\nabla}$  — дифференцирование по координатам ( $\hat{p}_x = -i\hbar\partial/\partial x$ ,  $\hat{p}_y = -i\hbar\partial/\partial y$ ,  $\hat{p}_z = -i\hbar\partial/\partial z$ ). ♦

При этом энергии частицы  $E = \vec{p}^2/2m + U(\vec{r})$  в потенциальном поле соответствует оператор

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta + U(\vec{r}). \quad (36.13)$$

Здесь оператор потенциальной энергии  $U(\vec{r})$  — просто умножение на функцию  $U(\vec{r})$ . В этих обозначениях уравнение Шрёдингера (36.11) имеет вид

$$i\hbar\frac{\partial}{\partial t}\Psi(t, \vec{r}) = \hat{H}\Psi(t, \vec{r}). \quad (36.14)$$

Оператор полной энергии  $\hat{H}$  называется *гамильтонианом*.

**Очень важно!** ♦ В классической механике законы сохранения связаны с симметрией системы. Энергия — с однородностью времени, т. е. с симметрией относительно его трансляции (сдвига)  $t \rightarrow t + \Delta t$ . Импульс — с однородностью пространства, т. е. с симметрией относительно трансляции пространства  $\vec{r} \rightarrow \vec{r} + \Delta\vec{r}$ . Момент импульса (количества движения) — с симметрией относительно поворотов (трансляции углов  $\varphi \rightarrow \varphi + \Delta\varphi$ ), т. е. с изотропией пространства. ♦

Трансляцию какой-то обобщенной координаты  $q$  производит оператор дифференцирования по этой координате. Например, для бесконечно малой трансляции  $q \rightarrow q + \Delta q$  имеем по определению производной

$$\Psi(q + \Delta q) \approx \Psi(q) + \Delta q \frac{\partial}{\partial q} \Psi(q).$$

Поэтому неслучайно в квантовой механике полной энергии соответствует  $\partial/\partial t$ , а импульсу — градиент  $\vec{\nabla}$ . Аналогично оператор проекции момента количества движения на какую-то ось  $z$  пропорционален оператору дифференцирования  $\partial/\partial\varphi$  по углу поворота вокруг этой оси:

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial\varphi}. \quad (36.15)$$

## 36.4 Стационарное уравнение Шрёдингера

В теории операторов важную роль играют так называемые *собственные состояния* операторов. Это такие состояния, которые при действии оператора меняются тривиальным образом: умножаются на некоторое число. Это число называется *собственным значением* данного оператора, соответствующим данному собственному состоянию. Чтобы найти собственные состояния и собственные значения какого-то оператора  $\hat{A}$  надо решить уравнение

$$\hat{A}\Psi_n(t, \vec{r}) = A_n\Psi_n(t, \vec{r}),$$

где индекс  $n$  отличает одно решение от другого. Набор величин  $A_n$ , т. е. набор собственных значений оператора  $\hat{A}$ , определяет его свойства.

Пример: операция поворота вокруг некоторой оси  $z$ . Роль состояний играют здесь обычные радиусы-векторы. Очевидно, что при повороте все векторы меняются, кроме параллельных оси. Это и есть собственные векторы оператора поворота вокруг оси  $z$ , причем соответствующее собственное значение равно единице. Аналогичны выводы для поворота вокруг осей  $x$  и  $y$ . Умножая какой-то из собственных векторов на число, мы снова получаем собственный вектор того же оператора поворота с тем же собственным значением. Это позволяет нам найти **нормированный** собственный вектор, длина которого равна единице. Произвольный поворот можно получить комбинацией этих трех поворотов. Соответственно любой радиус-вектор  $\vec{r}$  можно представить как линейную комбинацию трех собственных ортонормированных векторов:  $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$ .

Ситуация с квантово-механическими операторами по сути ничем не отличается от описанной: зная набор собственных состояний  $\Psi_n(t, \vec{r})$ , любое другое состояние  $\Psi(t, \vec{r})$  можно получить с помощью линейной комбинации, т. е. с помощью принципа суперпозиции<sup>2</sup>:

$$\Psi(t, \vec{r}) = \sum_n C_n \Psi_n(t, \vec{r}). \quad (36.16)$$

Связь математики с физикой реализуется в следующем правиле.

**Правило 2.**  $\blacklozenge$  *Единичное измерение любой физической величины  $A$  всегда дает лишь одно из собственных значений  $A_n$  соответствующего ей оператора  $\hat{A}$ . При повторе измерения результатом, вообще говоря, могут быть любые другие собственные значения  $A_m$ . Вероятность получить при измерении данное конкретное значение  $A_n$  определяется состоянием системы (а именно квадратом модуля  $|C_n|^2$  соответствующего коэффициента в разложении (36.16)).*  $\blacklozenge$

*Следствие:* в собственном состоянии  $\Psi_m(t, \vec{r})$  измерение  $A$  с вероятностью 100% даст значение  $A_m$  (так как в этом случае в разложении (36.16) отличен от нуля лишь коэффициент  $C_m = 1$ ).

Поскольку среди всех физических величин особую роль выполняет энергия, найдем уравнение для собственных состояний  $\Psi_E$  оператора полной энергии. Уравнение, согласно сказанному, имеет вид:

$$i\hbar \frac{\partial}{\partial t} \Psi_E(t, \vec{r}) = E \Psi_E(t, \vec{r}),$$

откуда следует решение

$$\boxed{\Psi_E(t, \vec{r}) = e^{-\frac{i}{\hbar} E t} \psi(\vec{r})}. \quad (36.17)$$

Мы получили общий вид состояния, в котором энергия имеет определенное значение. Такие состояния называются **стационарными**. Естественно, пока

<sup>2</sup>Индекс, отличающий одно собственное значение от другого, может быть и непрерывной величиной, а вместе с ним будут непрерывными и соответствующие собственные значения оператора  $\hat{A}$ . В этом случае суммирование в (36.16) заменяется на интегрирование.



невозможно сказать, чему равна энергия стационарного состояния, поскольку мы еще не указали рассматриваемую физическую систему. В уравнении (36.17) стоит некая функция  $\psi(\vec{r})$ , не зависящая уже от времени. Она называется **волновой функцией стационарного состояния**. Зависимость стационарных состояний от времени особенно проста — как для свободной частицы. Отсюда следует, что

$$|\Psi_E(t, \vec{r})|^2 = |\psi(\vec{r})|^2,$$

т.е. в стационарном состоянии плотность вероятности не зависит от времени. В этом смысле и следует понимать название «стационарное». Подставляя решение (36.17) в общее уравнение Шрёдингера (36.11), получаем **стационарное уравнение Шрёдингера**, т.е. уравнение для  $\psi(\vec{r})$ :

$$\boxed{-\frac{\hbar^2}{2m} \Delta \psi(\vec{r}) + U(\vec{r})\psi(\vec{r}) = E \psi(\vec{r})}. \quad (36.18)$$

Подчеркнем: это — уравнение для состояний с *определенной энергией*  $E$ . В операторных обозначениях оно имеет вид  $\hat{H}\psi(\vec{r}) = E\psi(\vec{r})$ , т.е. представляет собой уравнение для собственных состояний гамильтониана. Задавая тот или иной вид потенциальной энергии, мы конкретизируем систему и получаем стационарное уравнение Шрёдингера, решения которого и описывают квантовые свойства системы.

Не следует думать, что система может быть только в стационарном состоянии. Возьмем характерный пример: пусть  $\psi_1(\vec{r})$  и  $\psi_2(\vec{r})$  — два неких стационарных состояния какой-то системы с разными энергиями  $E_1$  и  $E_2$ . Предположим, что в начальный момент времени волновая функция системы является симметричной суперпозицией этих состояний:

$$\Psi(0, \vec{r}) = \psi_1(\vec{r}) + \psi_2(\vec{r}).$$

Вопрос: что будет с системой в произвольный момент  $t$ ?

Зная, что справедлив принцип суперпозиции и что зависимость собственных состояний от времени определяется соотношениями типа (36.17), можно сразу же написать волновую функцию:

$$\Psi(t, \vec{r}) = e^{-\frac{i}{\hbar}E_1 t} \psi_1(\vec{r}) + e^{-\frac{i}{\hbar}E_2 t} \psi_2(\vec{r}). \quad (36.19)$$

Плотность вероятности такого состояния зависит от времени! Введем обозначения для средней энергии  $E = (E_1 + E_2)/2$  и частоты перехода  $\omega = (E_1 - E_2)/\hbar$ . Тогда  $E_1 = E + \hbar\omega/2$ ,  $E_2 = E - \hbar\omega/2$ , и легко получаем<sup>3</sup> вместо (36.19):

$$\begin{aligned} \Psi(t, \vec{r}) &= e^{-\frac{i}{\hbar}Et} \left[ e^{-i\frac{\omega t}{2}} \psi_1(\vec{r}) + e^{i\frac{\omega t}{2}} \psi_2(\vec{r}) \right] = \\ &= e^{-\frac{i}{\hbar}Et} \left[ (\psi_1 + \psi_2) \cos \frac{\omega t}{2} - i(\psi_1 - \psi_2) \sin \frac{\omega t}{2} \right]. \end{aligned} \quad (36.20)$$

Видно, что в момент  $t = 0$  система находится в симметричном состоянии, к моменту времени  $t = \pi/\omega$  она перейдет в антисимметричное состояние, а

<sup>3</sup>Здесь применяется формула Эйлера  $e^{i\alpha} = \cos \alpha + i \sin \alpha$ , где  $i$  — мнимая единица.

в момент  $t = 2\pi/\omega$  снова вернется в симметричное состояние. Следовательно, система осциллирует между симметричным и антисимметричным состояниями с круговой частотой  $\omega$ . Здесь усматривается аналогия с классической физикой: в рассмотренной ранее системе связанных осцилляторов (см. разд. 24.4) возникают похожие собственные колебания (нормальные моды) и биения.

## 36.5 Уравнение Шрёдингера для простейших систем

### Свободная частица, движущаяся вдоль оси $x$

Потенциальная энергия равна нулю ( $U(x) = 0$ ), и производные по  $y$  и  $z$  в операторе Лапласа исчезают. Стационарное уравнение Шрёдингера (36.18) принимает вид

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x) = E\psi(x).$$

Введем волновое число  $k$ , связанное с энергией  $E$  частицы соотношением  $E = \hbar^2 k^2 / (2m)$ , и перепишем уравнение в виде

$$\psi''(x) + k^2 \psi(x) = 0. \quad (36.21)$$

Существуют, как известно, два линейно независимых решения дифференциального уравнения (36.21), так что общее решение можно представить как суперпозицию

$$Ae^{ikx} + Be^{-ikx}$$

(с учетом зависимости от времени (см. (36.5)) первый член соответствует волне, бегущей направо, вдоль положительного направления оси  $Ox$ ; второй описывает волну, бегущую налево, в противоположном направлении; постоянные  $A$  и  $B$  произвольны). Это же решение можно представить в виде пространственной части суперпозиции двух стоячих волн:

$$A \sin(kx) + B \cos(kx).$$

Аналогия: такие же решения описывают колебания свободной струны. Поскольку возможны волны с любой произвольной частотой, струна не звучит (это эквивалентно тому, что энергия частицы не квантуется). Для частицы, движущейся в произвольном направлении вдоль волнового вектора  $\vec{k}$ , справедливы те же самые решения при замене  $kx$  на  $\vec{k} \cdot \vec{r}$ .

При решении большинства задач квантовой механики следует обратить внимание на то, что волновая функция должна быть непрерывной — вероятность пребывания частицы не может меняться скачком от точки к точке.

А если потенциал конечен, то из уравнения Шрёдингера следует, что первая производная также непрерывна. Действительно, проинтегрируем уравнение Шрёдингера по бесконечно малой окрестности  $[a - \epsilon, a + \epsilon]$ , охватывающей некоторую точку  $a$ :

$$-\frac{\hbar^2}{2m} \int_{a-\epsilon}^{a+\epsilon} \psi''(x) dx + \int_{a-\epsilon}^{a+\epsilon} U(x) \psi(x) dx = E \int_{a-\epsilon}^{a+\epsilon} \psi(x) dx. \quad (36.22)$$

Первое слагаемое в левой части (36.22) дает скачок производной в точке  $a$ , а в силу непрерывности волновой функции можно вынести ее значение в точке  $a$  из-под знака двух других интегралов:

$$-\frac{\hbar^2}{2m} [\psi'(a + \epsilon) - \psi'(a - \epsilon)] + \psi(a) \int_{a-\epsilon}^{a+\epsilon} U(x) dx = E\psi(a) \int_{a-\epsilon}^{a+\epsilon} dx. \quad (36.23)$$

Правая часть равна нулю в силу бесконечной малости интервала интегрирования. Второе слагаемое в левой части (36.23) равно нулю по этой же причине, но при условии, что потенциал на этом интервале конечен (сам потенциал или его производная не обязаны быть непрерывными). Таким образом, мы приходим к выводу о непрерывности первой производной волновой функции:

$$\psi'(a + \epsilon) = \psi'(a - \epsilon) \quad (36.24)$$

### Частица в бесконечно глубокой потенциальной яме

Потенциал в этой задаче имеет вид

$$U(x) = \begin{cases} 0, & \text{если } 0 \leq x \leq l, \\ +\infty, & \text{если } x > l \text{ или } x < 0. \end{cases} \quad (36.25)$$

Такая система соответствует частице, движущейся вдоль прямой линии и отскакивающей от абсолютно отражающих препятствий в точках  $x = 0$  и  $x = l$ . В область бесконечного потенциала частица проникнуть не может, следовательно  $\psi(x) = 0$  за пределами отрезка  $[0, l]$ . Внутри ямы  $U(x) = 0$ , и стационарное уравнение Шрёдингера имеет тот же вид, что и для свободной частицы. Получатся те же решения в виде суперпозиции стоячих (или бегущих) волн, но, в отличие от предыдущего случая, добавятся граничные условия. Именно в точках  $x = 0$  и  $x = l$  волновая функция должна обращаться в нуль (поскольку она непрерывна и равна нулю вне ямы). В классической механике точно такие же граничные условия имеет уравнение для закрепленной струны.

Общее решение берем в виде  $\psi(x) = A \sin(kx) + B \cos(kx)$ . Используем сначала первое граничное условие:

$$\psi(0) = B = 0.$$

Мы получили, что решение уравнения Шрёдингера должно иметь вид  $\psi(x) = A \sin(kx)$ . Если продолжить нашу аналогию, то можно сказать, что на струне, закрепленной в одной точке, бегущих волн не бывает: отражение от неподвижной точки обязательно порождает стоячую волну. Однако на длину волны никаких ограничений не накладывается: такая струна также не звучит.

Теперь используем второе из граничных условий:

$$\psi(l) = A \sin(kl) = 0.$$

Здесь есть два типа решений. При  $A = 0$  получаем  $\psi(x) \equiv 0$ , что означает отсутствие частицы в яме (вероятность найти ее всюду равна нулю). Поэтому нас интересует нетривиальное второе решение, когда  $\sin(kl) = 0$ . Это возможно

лишь при некоторых значениях  $k_n$  волнового числа, когда  $k_n l = \pi(n + 1)$ , где  $n = 0, 1, 2, \dots$ . Так как энергия частицы связана с волновым числом, то

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\pi^2 \hbar^2 (n + 1)^2}{2ml^2}, \quad n = 0, 1, 2, \dots \quad (36.26)$$

Мы получили квантование энергии, т. е. наша «струна», закрепленная с обеих сторон, «звучала», так как появились выделенные частоты.

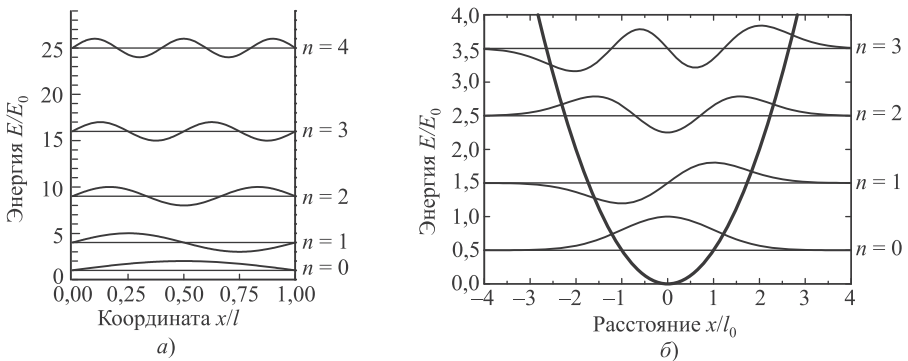
Подставляя найденные разрешенные значения волнового числа в выражение для волновой функции, получаем ее в виде

$$\psi_n(x) = \begin{cases} 0, & \text{если } x < 0 \text{ или } x > l; \\ \sqrt{\frac{2}{l}} \sin\left(\pi(n + 1)\frac{x}{l}\right), & \text{если } 0 \leq x \leq l. \end{cases} \quad (36.27)$$

Значение постоянной  $A = \sqrt{2/l}$  определено из условия нормировки (36.3).

Схема уровней энергии и вид волновых функций для первых пяти уровней показаны на рис. 36.1, а. **Квантовое число**  $n$ , нумерующее энергетические уровни, равно **числу узлов** (нулей) волновой функции. Так, волновая функция основного состояния ( $n = 0$ ) узлов не имеет, у волновой функции первого возбужденного состояния ( $n = 1$ ) один узел и т. д. Это наблюдение имеет общий характер.

Откуда же берется дискретность уровней энергии, характерная и для атома? Сравним со свободной частицей: уравнения те же, но с иными граничными условиями! Здесь возможны две постановки задачи. В первом случае исследуется состояние, которому в *классической механике* соответствовало бы **инфинитное движение** (задача рассеяния). Обычно в таких случаях решения возможны при *любых* значениях энергии  $E$  (как говорят, спектр *непрерывен*). Во втором случае исследуется состояние, которому в *классической механике* соответствует



**Рис. 36.1.** Схема уровней энергии в бесконечной прямоугольной яме (а) и для гармонического осциллятора (б); тонкими линиями показаны соответствующие волновые функции

**финитное движение** в ограниченной области пространства (задача на связанные состояния). Требование конечности волновой функции во всем пространстве ведет к **квантованию энергии**. Подчеркнем: в этом случае стационарное уравнение имеет физически приемлемые решения не всегда, а лишь при некоторых значениях  $E$ . Как следствие возникает **дискретный спектр энергии** системы.

**Задача 36.46.** Определить разность энергий соседних уровней  $\Delta E$  для частицы в бесконечно высокой потенциальной яме при больших значениях  $n$ . Полученный результат использовать для оценки разности энергий соседних уровней молекул азота при комнатной температуре  $T = 300$  К в сосуде. Масса молекулы  $m = 2,3 \cdot 10^{-26}$  кг, а линейный размер сосуда  $l = 0,1$  м. Сравнить результат с кинетической энергией поступательного движения молекул азота.

**Решение.** Используя выражение (36.26) для уровней энергии частицы в потенциальной яме, находим разность энергий соседних уровней:

$$\Delta E = \frac{\hbar^2 \pi^2}{2ml^2} [(n+1)^2 - n^2] = \frac{\hbar^2 \pi^2}{2ml^2} (2n+1) \approx \frac{\hbar^2 \pi^2}{ml^2} n$$

при больших значениях  $n$ . Средняя кинетическая энергия поступательного движения молекул идеального газа равна  $E = 3k_B T/2 \approx 6,21 \cdot 10^{-21}$  Дж. Приравнявая  $E$  к выражению для энергии уровней частицы в яме, находим, что такая энергия соответствует квантовым числам порядка

$$n = \frac{l}{\pi \hbar} \sqrt{3mk_B T} \approx 5 \cdot 10^9.$$

Уже само по себе это число говорит, что мы находимся в области крайне высоких возбуждений, т. е. в области справедливости классических законов. Разность энергий соседних уровней получаем, подставляя в формулу для  $\Delta E$  найденное выражение для квантового числа  $n$  с учетом массы  $m$  молекул азота:

$$\Delta E = \frac{\pi \hbar}{l} \sqrt{\frac{3k_B T}{m}} \approx 2,4 \cdot 10^{-30} \text{ Дж.}$$

В электронвольтах те же характеристики имеют значения  $E \approx 0,039$  эВ,  $\Delta E \approx 1,5 \cdot 10^{-11}$  эВ. Относительная разность энергии уровней ничтожно мала:  $\Delta E/E \approx 4 \cdot 10^{-10}$ , и потому в классической физике квантовой дискретностью пренебрегают. ■

### Частица в трехмерной потенциальной яме

Это — обобщение предыдущей задачи. Частица может двигаться в кубическом потенциальном объеме с длиной ребра  $l$ . Нетрудно убедиться, что общее решение для волновой функции представимо в виде произведения одномерных волновых функций, полученных в предыдущей задаче:

$$\psi_{n_1 n_2 n_3}(x, y, z) = \left(\frac{2}{l}\right)^{\frac{3}{2}} \sin\left(\pi(n_1+1)\frac{x}{l}\right) \sin\left(\pi(n_2+1)\frac{y}{l}\right) \sin\left(\pi(n_3+1)\frac{z}{l}\right). \quad (36.28)$$

Такая волновая функция соответствует очевидному факту, что движения вдоль трех осей не зависят друг от друга и каждое описывается прежними одномерными волновыми функциями. Энергия, как легко догадаться, будет равна сумме энергий движения по осям  $x, y, z$ :

$$E_{n_1 n_2 n_3} = \frac{\hbar^2 \pi^2}{2ml^2} ((n_1+1)^2 + (n_2+1)^2 + (n_3+1)^2). \quad (36.29)$$

Состояние системы теперь определяется тремя квантовыми числами  $n_1$ ,  $n_2$  и  $n_3$ , принимающими, как и прежде, целые значения. Здесь мы впервые сталкиваемся с важным понятием **вырождения** энергетических уровней, т.е. с ситуацией, когда разные состояния системы имеют одну ту же энергию. В самом деле, минимальная энергия системы достигается при минимальных значениях всех квантовых чисел, т.е. при  $n_1 = n_2 = n_3 = 0$ . Эта энергия равна:

$$E_0 = \frac{3 \hbar^2 \pi^2}{2 ml^2},$$

и ей соответствует одна волновая функция  $\psi_{000}$ . Говорят, что **основное состояние не вырождено** (невыврожденность состояния с минимальной энергией — общее правило). Первое возбужденное состояние получается, когда *одно* из квантовых чисел равно единице, а остальные по-прежнему равны нулю; энергия этого возбужденного состояния равна:

$$E_1 = 3 \frac{\hbar^2 \pi^2}{ml^2}.$$

Но такая энергия соответствует теперь трем различным состояниям с волновыми функциями  $\psi_{001}$ ,  $\psi_{010}$ ,  $\psi_{100}$  (квантовое число 1 можно выбрать *тремя способами*). В таком случае говорят, что **кратность вырождения** первого возбужденного уровня равна трем. Естественно, в другой системе может быть совершенно иная кратность вырождения возбужденных уровней (или отсутствие такового). Последующие состояния (или другие, более высокие уровни энергии) для частицы в потенциальной яме с бесконечно высокими стенками также вырождены. Ясно, что вырождение уровней связано с симметрией системы, с равноправием всех осей. Если бы размеры ямы были разными ( $l_1$ ,  $l_2$ ,  $l_3$ ) по всем трем направлениям, то для энергии мы бы получили вместо (36.29) выражение

$$E_{n_1 n_2 n_3} = \frac{\hbar^2 \pi^2}{2m} \left( \frac{(n_1 + 1)^2}{l_1^2} + \frac{(n_2 + 1)^2}{l_2^2} + \frac{(n_3 + 1)^2}{l_3^2} \right)$$

и вырождение могло бы иметь место лишь случайным образом, при определенных соотношениях между длиной, шириной и высотой потенциального ящика.

## Одномерный осциллятор

В классической физике осциллятор представляет собой тело, прикрепленное к пружине и колеблющееся с круговой частотой  $\omega$ . Потенциальная энергия такой системы описывается выражением  $U(x) = m\omega^2 x^2/2$ , так что уравнение Шрёдингера можно записать в виде

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x) + \frac{m\omega^2}{2} x^2 \psi(x) = E \psi(x).$$

Отсюда можно найти решение для волновой функции основного состояния:

$$\psi(x) = \sqrt[4]{\frac{m\omega}{\pi\hbar}} \exp\left(-\frac{m\omega x^2}{2\hbar}\right).$$

Она существенно отличается от нуля в области с характерным размером  $l_0 = \sqrt{\hbar/(m\omega)}$ . Подставляя это выражение в уравнение Шрёдингера, можем убедиться, что энергия основного состояния равна  $E_0 = \hbar\omega/2$ . Мы не выписываем волновые функции возбужденных состояний осциллятора, но выражение для разрешенных значений энергии имеет вид ( $n$  — колебательное квантовое число)

$$E_n = \hbar\omega \left( n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad (36.30)$$

Здесь воспроизводятся формула Планка и нулевые колебания  $E_0 = \hbar\omega/2$ , полученные ранее из соотношения неопределенностей (см. разд. 35.3).

Схема уровней энергии и вид волновых функций для первых четырех уровней показаны на рис. 36.1, б. Как и в потенциальной яме с бесконечно высокими стенками, здесь квантовое число  $n$  тоже равно числу узлов волновой функции.

Осциллятор играет совершенно особую роль в физике. В квантовой электродинамике с его помощью описываются квантованные колебания электромагнитного поля — фотоны. Но и в нерелятивистской квантовой механике его роль велика. Действительно, пусть задан произвольный потенциал  $U(x)$ , в котором есть точка, где потенциальная энергия достигает минимума. Поместим туда начало координат:  $x = 0$ . Точка минимума потенциала — это точка равновесия: здесь производная потенциала равна нулю, т. е. равна нулю сила  $F = -dU/dx$ , действующая на частицу. Если равновесие устойчивое, то вторая производная потенциала положительна:

$$\left. \frac{dU(x)}{dx} \right|_{x=0} = 0, \quad \left. \frac{d^2U(x)}{dx^2} \right|_{x=0} = k > 0.$$

При малых отклонениях координаты от точки минимума потенциальную энергию можно разложить в ряд Тейлора:

$$U(x) = U(0) + x \left( \left. \frac{dU(x)}{dx} \right|_{x=0} \right) + \frac{x^2}{2} \left( \left. \frac{d^2U(x)}{dx^2} \right|_{x=0} \right) + \dots \approx U(0) + \frac{kx^2}{2} + \dots$$

Таким образом, в классической физике малые колебания вблизи положения устойчивого равновесия всегда гармонические. А в квантовой механике низколежащие уровни энергии (в частности, нулевые колебания) могут быть описаны в терминах квантового осциллятора (см. разд. 35.3).

### Трехмерный осциллятор

Эта задача является обобщением предыдущей. Как и для трехмерной потенциальной ямы с бесконечно высокими стенками, волновая функция представляется в виде произведения волновых функций одномерных осцилляторов, колеблющихся независимо вдоль осей  $x, y, z$ . Так, волновая функция основного состояния имеет вид

$$\psi_{000}(\vec{r}) = \left( \frac{m\omega}{\pi\hbar} \right)^{\frac{3}{2}} \exp \left( -\frac{m\omega\vec{r}^2}{2\hbar} \right),$$

а уровни энергии трехмерного осциллятора описываются формулой

$$E_{n_1 n_2 n_3} = \hbar\omega \left( n_1 + n_2 + n_3 + \frac{3}{2} \right), \quad n_i = 0, 1, 2, \dots \quad (36.31)$$

В отличие от одномерного осциллятора состояние определяется значениями трех квантовых чисел  $n_1, n_2, n_3$ . Легко понять, что все возбужденные состояния должны быть вырожденными.

## 36.6 Принцип соответствия Бора

В методологии науки принцип соответствия — утверждение, что любая новая научная теория при наличии старой, хорошо проверенной теории находится с ней не в полном противоречии, а дает те же следствия в некотором предельном приближении (частном случае). Например, в специальной теории относительности (СТО) в пределе малых скоростей  $v \ll c$  получаются те же следствия, что и в классической механике: преобразования Лоренца переходят в преобразования Галилея, время течет одинаково во всех системах отсчета, кинетическая энергия становится равной  $mv^2/2$  и т. д.

В квантовой механике *принципом соответствия* называется утверждение о том, что квантово-механическое описание системы переходит в классическое в пределе больших квантовых чисел (Н. Бор, 1923 г.). Обычные в нашем мире значения энергии велики по сравнению с характерной энергией основного состояния и расщеплением уровней: с высокой лестницы мы не различаем ступенек. То же самое скажем более научно: *при больших квантовых числах* (высоколежащих уровнях) должны воспроизводиться *классические результаты*. Покажем это на примере атома водорода.

В разделе 35.1 было получено классическое выражение для скорости электрона на орбите радиусом  $r$ :

$$v = e\sqrt{\frac{Z}{4\pi\epsilon_0 m_e r}}.$$

Отсюда легко найти классическую частоту вращения электрона:

$$\nu_{cl} = \frac{v}{2\pi r} = \frac{e}{2\pi r} \sqrt{\frac{Z}{4\pi\epsilon_0 m_e r}}.$$

Классическое выражение для энергии электрона на орбите

$$E = -\frac{Ze^2}{8\pi\epsilon_0 r}$$

позволяет выразить радиус орбиты через энергию электрона:

$$r = \frac{Ze^2}{8\pi\epsilon_0 |E|}.$$



Подставляя это выражение в формулу для классической частоты вращения  $\nu_{cl}$ , получаем

$$\nu_{cl} = \frac{4\varepsilon_0}{Ze^2} \sqrt{\frac{2|E|^3}{m_e}}. \quad (36.32)$$

Именно на этой частоте ожидается излучение электрона в классической теории.

В разделе 35.1 было также выведено выражение для энергии уровня с номером  $n$

$$E_n = -\frac{m_e Z^2 e^4}{8(\varepsilon_0 h)^2 n^2}.$$

При  $n \gg 1$  получаем отсюда квантовую частоту перехода  $\nu_q$  между соседними уровнями:

$$\nu_q = \frac{E_n - E_{n-1}}{h} = \frac{m_e Z^2 e^4}{8(\varepsilon_0 h)^2} \left( \frac{1}{(n-1)^2} - \frac{1}{n^2} \right) \approx \frac{m_e Z^2 e^4}{8(\varepsilon_0 h)^2} \frac{2}{n^3}.$$

Выражая квантовое число  $n$  через энергию уровня, находим:

$$n = \frac{Ze^2}{2\varepsilon_0 h} \sqrt{\frac{m_e}{2|E_n|}}.$$

Подставляя это выражение в формулу для квантовой частоты перехода между соседними высоколежащими уровнями, приходим к окончательному результату

$$\nu_q = \frac{4\varepsilon_0}{Ze^2} \sqrt{\frac{2|E_n|^3}{m_e}}. \quad (36.33)$$

На этой частоте должен излучать сильно возбужденный атом Бора. Сравнивая классическую частоту (36.32) с квантовой (36.33), убеждаемся, что при  $E = |E_n|$  они совпадают. Это свойственно не только водородоподобному атому. Аналогичный результат получается для бесконечной потенциальной ямы, этот же вывод можно сделать и для прочих систем. Следовательно, соблюдается **принцип соответствия Бора**:

◆ *классическая механика является предельным случаем квантовой.* ◆

**Задача 36.47.** Используя формулы и данные задачи 36.46, продемонстрировать справедливость принципа соответствия Бора для молекул азота в сосуде.

**Решение.** При переходе молекул между уровнями с разностью энергий  $\Delta E$  излучается квант света с энергией  $h\nu = \Delta E$ , откуда находим

$$\nu = \frac{1}{2l} \sqrt{\frac{3k_B T}{m}}.$$

Кроме того, классическая скорость поступательного движения молекул азота равна  $v = \sqrt{3k_B T/m}$ , и они пролетают в сосуде от стенки до стенки и обратно за время  $\tau = 2l/v$  (период классического движения). Обратная величина есть классическая частота

$$\nu_{cl} = \frac{v}{2l} = \frac{1}{2l} \sqrt{\frac{3k_B T}{m}}.$$

Именно на этой частоте классическая физика предсказывает электромагнитное излучение. Принцип соответствия Бора заключается в том, что совпадают обе частоты, квантовая — при переходах между высоковозбужденными состояниями — и классическая. (То же самое справедливо, как мы видели, для атома Бора.) Если подставить численные значения, то для частоты излучения в данной задаче получится величина  $\nu \approx 3,7$  кГц, что соответствует длине волны  $\lambda = c/\nu \approx 80$  км, т. е. сверхдлинным радиоволнам. ■

Квантовая механика описывает поведение микрочастиц, классическая механика хороша для макрообъектов. Обе подтверждены на опыте. Однако законы физики не должны зависеть от размеров описываемых физических объектов. Природа едина, а классическая физика — это приближение к квантовой физике, когда рассматриваемые системы становятся большими. Но что значит большие? Бор ввел интуитивное определение — большие квантовые числа. Более современная формулировка говорит, что классическое приближение справедливо при больших значениях *действия*  $S$ . Эта скалярная величина играет важную роль в теоретической механике, так как физические системы движутся по траекториям, на которых величина действия минимальна (*принцип наименьшего действия*). Но нам сейчас достаточно знать, что размерность действия — Дж·с, т. е. совпадает с размерностью постоянной Планка. Таким образом, для больших систем  $S \gg \hbar$ . Такую же размерность имеет произведение энергии на время или импульса на расстояние. Поэтому критерий «большой» системы может быть записан как  $pl \gg \hbar$  или  $Et \gg \hbar$ , т. е. произведение характерного импульса процесса на характерный размер системы и произведение характерной энергии процесса на его характерное время должны быть значительно больше постоянной Планка. В таком случае система может считаться большой и будет описываться законами классической физики.

## 36.7 Отражение и туннелирование частиц

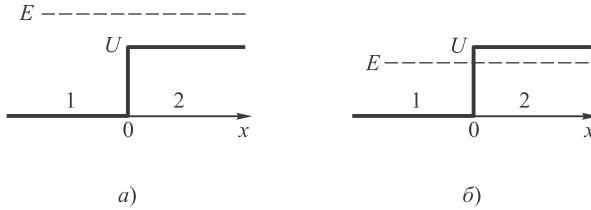
До сих пор мы имели дело с задачами на связанные состояния. Рассмотрим теперь примеры инфинитного движения частиц, когда они могут уходить на бесконечно большие расстояния. В простейшем случае движения вдоль одной из координатных осей задача рассеяния частиц сводится к задаче взаимодействия частицы с неким потенциальным барьером. Мы рассмотрим ряд барьеров простой прямоугольной формы, чтобы усвоить характерные особенности этого типа квантовых явлений.

### Низкий бесконечный барьер

Потенциальная энергия имеет вид

$$U(x) = \begin{cases} 0, & \text{при } x < 0 \quad (\text{область 1}), \\ U, & \text{при } x > 0 \quad (\text{область 2}). \end{cases} \quad (36.34)$$

Слово «низкий» означает, что высота барьера  $U$  меньше энергии частицы  $E$  (рис. 36.2, а). Решим уравнение Шрёдингера отдельно для каждой из областей. В области 1 потенциал равен нулю, и мы имеем то же общее решение, что и для



**Рис. 36.2.** Низкий (а) и высокий (б) потенциальные барьеры; штриховой линией показана энергия налетающей слева частицы, цифрами — номера областей с различной потенциальной энергией

свободной частицы:

$$\psi_1(x) = Ae^{ik_1x} + Be^{-ik_1x}, \quad E = \frac{\hbar^2 k_1^2}{2m},$$

где  $A$  и  $B$  — амплитуды падающей и отраженной волн соответственно. В области 2 уравнение Шрёдингера имеет вид:

$$-\frac{\hbar^2}{2m}\psi''(x) + U\psi(x) = E\psi(x).$$

В области 2 меняются кинетическая энергия и импульс частицы, и мы должны ввести другое волновое число  $k_2$  в отличие от прежнего  $k_1$ :

$$E - U = \frac{\hbar^2 k_2^2}{2m}.$$

Тогда очевидно, что решение уравнения Шрёдингера в области 2 будет иметь тот же вид, что и в области 1, но с заменой  $k_1$  на  $k_2$ . Однако из физических соображений ясно, что в области 2 не может быть волны, распространяющейся справа налево (в бесконечно удаленной точке ей не от чего отражаться). Поэтому волновая функция в этой области соответствует прямой волне (бегущей вправо, по нашей терминологии):

$$\psi_2(x) = Ce^{ik_2x}.$$

По сути, здесь мы снова использовали некое граничное условие (отсутствие отраженной волны), хотя и иное, нежели для задачи на связанные состояния. Нам осталось определить только амплитуды  $A$ ,  $B$ ,  $C$ .

Для этого мы должны вспомнить, что  $\psi_1(x)$  и  $\psi_2(x)$  — значения одной волновой функции в разных пространственных областях. Эта волновая функция должна быть непрерывна вместе со своей первой производной по переменной  $x$ . Непрерывность функции в точке  $x = 0$  означает, что должно выполняться условие

$$\psi_1(0) = \psi_2(0), \quad \text{откуда} \quad A + B = C.$$

Непрерывность первой производной волновой функции означает выполнение равенства

$$\psi_1'(0) = \psi_2'(0), \quad \text{откуда} \quad k_1(A - B) = k_2C.$$

Решение двух полученных уравнений дает

$$C = A \frac{2k_1}{k_1 + k_2}, \quad B = A \frac{k_1 - k_2}{k_1 + k_2}.$$

Амплитуда падающей волны остается неопределенной: ясно, что она зависит от интенсивности потока частиц. Важны не сами амплитуды, а отношение  $R$  квадратов их модулей, т. е. интенсивностей отраженной и падающей волн:

$$R = \frac{|B|^2}{|A|^2} = \left| \frac{k_1 - k_2}{k_1 + k_2} \right|^2. \quad (36.35)$$

Величина  $R$  называется **коэффициентом отражения**. По физическому смыслу это вероятность отражения частицы от барьера. Соответственно величина  $D = 1 - R$ , называемая **коэффициентом прохождения**, определяет вероятность проникновения частицы в правую область (за барьером). При действительных волновых числах  $k_1$  и  $k_2$  из (36.35) получаем для  $D$  выражение

$$D = \frac{4k_1 k_2}{(k_1 + k_2)^2}, \quad (36.36)$$

Удивительно, что частица имеет шанс отразиться от низкого барьера и повернуть назад. В классической физике частица не отразится от барьера ( $R = 0$ ), если ей хватит энергии пройти над ним (т. е. при  $E > \max U$ ). Например, с точки зрения классической физики электрон с энергией 10 эВ, влетевший в конденсатор с тормозящим потенциалом 5 В, безусловно, преодолет торможение и продолжит свой путь с уменьшенной энергией 5 эВ. В квантовой же теории не равна нулю вероятность того, что электрон отразится от поля конденсатора и повернет назад.

### Высокий бесконечный барьер

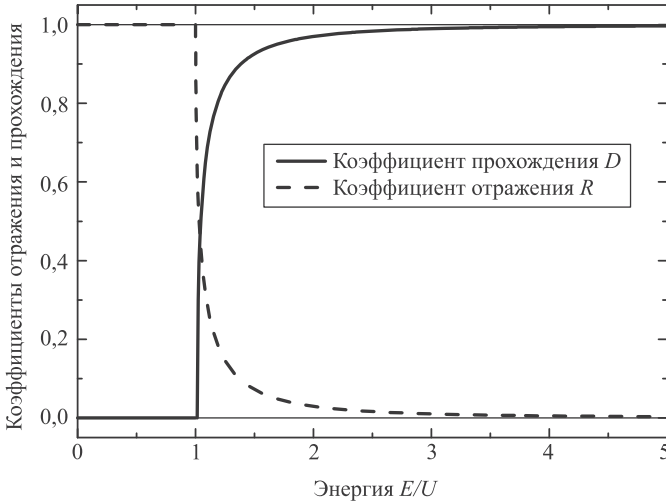
Потенциальная энергия имеет тот же вид, но энергия частицы меньше высоты барьера:  $E < U$  (рис. 36.2, б). Решение в области 1 остается прежним: суперпозиция прямой и отраженной волн. В области же 2 из-за обратного соотношения между энергией частицы и высотой барьера волновой вектор становится мнимым:

$$k_2 = i\kappa, \quad \text{где } \kappa = \frac{1}{\hbar} \sqrt{2m(U - E)}.$$

При подстановке мнимого волнового числа  $k_2 = i\kappa$  в выражение (36.35) для коэффициента отражения  $R$  получаем, что  $R = 1$ . Как и в классике, частица с энергией, меньшей высоты бесконечного барьера, наверняка отразится от него. Правда, в классической физике частица вовсе не может проникнуть под барьер. Наше же решение уравнения Шрёдингера для области 2 в случае высокого барьера становится равным

$$\psi_2(x) = C e^{-\kappa x}.$$

Это уже не совсем волна, а экспоненциально затухающая функция. Как и в случае низкого барьера, отброшено нефизическое решение — экспоненциально



**Рис. 36.3.** Коэффициенты отражения и прохождения в зависимости от отношения энергии частицы к высоте потенциального барьера; область  $E/U < 1$  соответствует высокому барьеру, область  $E/U > 1$  — низкому

растущая функция вида  $e^{\kappa x}$ . Под глубиной проникновения частицы под барьер  $d$  принято понимать расстояние, на котором интенсивность потока (вероятность) ослабевает в  $e$  раз. Из выражения для  $\psi_2(x)$  следует, что  $d = 1/(2\kappa)$ .

Зависимость коэффициентов отражения и прохождения от отношения  $E/U$  энергии частицы к высоте потенциального барьера показана на рис. 36.3.

### Конечный потенциальный барьер

Представим себе ледяную горку, по которой без трения может скользить хоккейная шайба (рис. 36.4). Ось  $Ox$  направим по горизонтали, а ось  $Oz$  — по вертикали. Пусть профиль горки описывается функцией  $z(x)$ , так что потенциальная энергия шайбы на горке равна  $U(x) = mgz(x)$ . Поскольку полная энергия  $E$  шайбы сохраняется, ее скорость  $v(x)$  уменьшается по мере подъема на горку:

$$\frac{mv^2(x)}{2} + U(x) = E.$$

**Точкой поворота**  $x_E$  называется место, где шайба останавливается и начинает движение назад. В этой точке кинетическая энергия шайбы равна нулю, так что  $U(x_E) = E$ , а высота подъема шайбы  $z(x_E) = E/(mg)$ . Чем больше полная энергия шайбы (ее начальная скорость  $v(0)$ ), тем выше заберется шайба на горку. Перевалить через нее она может лишь при энергии  $E > U_{\max}$ . При меньших энергиях шайба никогда не окажется по другую сторону горки. Для «квантовой шайбы» такое возможно.

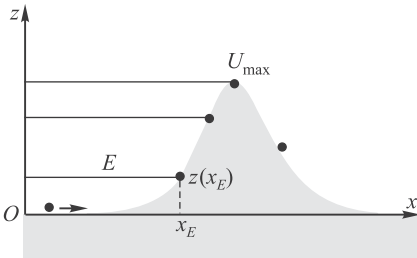


Рис. 36.4. Шайба на ледяной горке

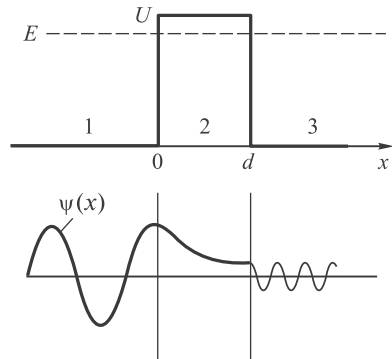


Рис. 36.5. Квантовая частица способна туннелировать сквозь конечный потенциальный барьер

Для простоты моделируем нашу «квантовую горку» прямоугольным потенциалом шириной  $d$  и высотой  $U$ :

$$U(x) = \begin{cases} 0, & \text{при } x < 0 & (\text{область 1}), \\ U, & \text{при } 0 \leq x \leq d & (\text{область 2}), \\ 0, & \text{при } x > d & (\text{область 3}). \end{cases} \quad (36.37)$$

Ясно, что происходит при  $E > U$ : с некоей вероятностью частица может отразиться от барьера. Этот процесс мы рассмотрим в следующем разделе. Но сейчас нас интересует случай  $E < U$ .

Мы видели, что интенсивность (квадрат модуля амплитуды) волны убывает под барьером и на расстоянии  $d$  становится меньше в  $\exp(-2d\kappa)$  раз. Но в этой точке барьер кончается, так что волна выйдет на свободу с уменьшенной амплитудой (рис. 36.5). Отношение интенсивностей выходящей и падающей волн называется **коэффициентом прозрачности** (эта величина также равна вероятности прохождения через барьер). Из наших рассуждений следует приближенное выражение для коэффициента прозрачности

$$D \approx \exp(-2\kappa d) = \exp\left[-\frac{2}{\hbar}d\sqrt{2m(U-E)}\right]. \quad (36.38)$$

Получая  $D$ , мы опустили некие множители перед экспонентой, что по физическому смыслу означает пренебрежение процессами, когда частица, прежде чем выйти из-под барьера, испытает многократное отражение от его стенок. При высоком и широком барьере ( $D \ll 1$ ) вклад таких процессов невелик и сделанное приближение оправдано.

Проникновение частицы сквозь конечный потенциальный барьер возможно в квантовой механике, но категорически запрещено в классической. В самом деле, формально вектор  $\hbar\kappa$  играет роль импульса (мнимого), так что кинетическая энергия  $T = -\hbar^2\kappa^2/(2m)$  становится отрицательной. Дело спасают соотношения неопределенностей. Модуль скорости (мнимой) частицы имеет порядок

$v \approx \hbar k/m$ , так что время туннелирования  $\Delta t \approx md/(\hbar \kappa)$ . Неопределенность в кинетической энергии  $\Delta T \approx \hbar/(2\Delta t) \approx \hbar^2 \kappa/(2md)$ . Из полученных результатов для коэффициента прозрачности видно, что эффект туннелирования заметен, если  $\kappa d < 1$ . Но тогда

$$\Delta T > \frac{\hbar^2 \kappa^2}{2m} = U - E.$$

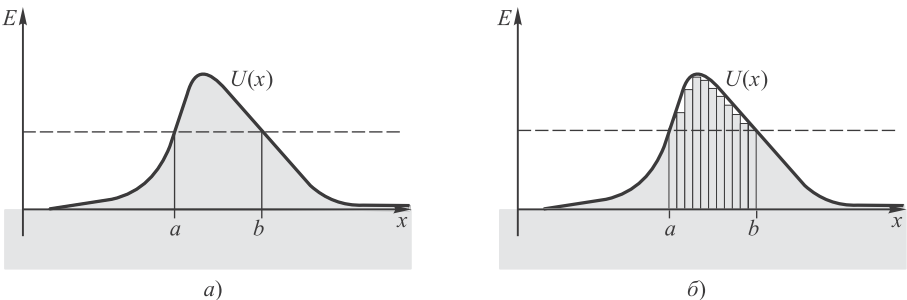
Получается, что неопределенность в кинетической энергии частицы под барьером больше самого значения кинетической энергии. Поэтому нельзя утверждать, что под барьером кинетическая энергия отрицательна. Скорее, она «размыта» настолько, что частица может как бы перепрыгнуть не слишком большой барьер. В случае же высокого и широкого барьера «размытость» кинетической энергии должна быть очень велика, что возможно лишь на очень короткое время, за которое частица не успевает проскочить за барьер. Поэтому в этом случае коэффициент прозрачности становится экспоненциально малым. По-другому: туннелирование заметно при ширине барьера порядка длины волны де Бройля.

Барьер произвольной формы (рис. 36.6) можно представить в виде последовательности прямоугольных барьеров равной ширины  $\Delta x$ , расположенных в точках  $x_i$  и имеющих высоту  $U(x_i)$ . Вероятность туннелирования через каждый из них мы знаем:

$$D_i \approx \exp \left[ -\frac{2}{\hbar} \Delta x \sqrt{2m(U(x_i) - E)} \right]. \quad (36.39)$$

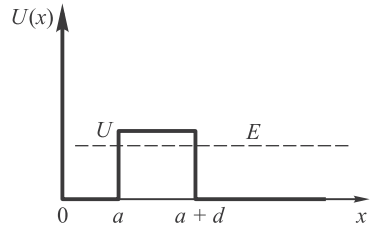
Вероятность туннелирования через всю совокупность прямоугольных барьеров выражается произведением вероятностей туннелирования через каждый из них:

$$D = \prod_i D_i \approx \exp \left[ -\frac{2}{\hbar} \sum_i \Delta x \sqrt{2m(U(x_i) - E)} \right]. \quad (36.40)$$



**Рис. 36.6.** Туннелирование через барьер произвольной формы (а) можно представить как последовательное туннелирование через прямоугольные барьеры бесконечно малой ширины (б)

**Рис. 36.7.** Частица в потенциальной яме, образованной непроницаемым препятствием слева и конечным барьером справа (к задаче 36.48)



Переход в (36.40) к пределу  $\Delta x \rightarrow 0$  ведет к появлению интеграла в экспоненте, так что вместо (36.38) имеем

$$D \approx \exp \left[ -\frac{2}{\hbar} \int_a^b dx \sqrt{2m(U(x) - E)} \right]. \quad (36.41)$$

Интеграл берется между точками поворота  $a$  и  $b$ , в которых потенциальная энергия равна полной ( $U(a) = U(b) = E$ ) и где классическая частица должна изменить направление движения.

**Задача 36.48.** Электрон находится в одномерной потенциальной яме шириной  $a = 10^{-10}$  м (рис. 36.7) и имеет энергию  $E = 1,5$  эВ. С одной стороны ямы потенциальная энергия  $U(x)$  бесконечна, а с другой стороны выйти из ямы электрону мешает потенциальный барьер высотой  $U = 2$  эВ и шириной  $d = 3 \cdot 10^{-10}$  м. Оценить время жизни  $\tau$  электрона в яме.

**Решение.** Скорость электрона в яме  $v = \sqrt{2E/m}$ , и за промежуток времени  $t$  он подойдет к барьеру  $vt/(2a)$  раз. При каждом подходе вероятность туннелирования равна  $D$ , так что вероятность туннелирования  $W_t$  за время  $t$  равна:

$$W_t \approx D \frac{vt}{2a} = t \frac{D}{a} \sqrt{\frac{E}{2m}}.$$

Вероятность  $W_t$  увеличивается с ростом промежутка времени  $t$ . При некотором значении  $t = \tau$  вероятность туннелирования станет равной единице и электрон вырвется из ямы. Отсюда получаем для времени жизни электрона в яме оценку:

$$\tau \approx \frac{a}{D} \sqrt{\frac{2m}{E}} \approx \frac{a}{c} \sqrt{\frac{2mc^2}{E}} \exp \left[ \frac{2d}{\hbar} \sqrt{2m(U - E)} \right].$$

Теперь остается подставить численные данные. Для упрощения вычислений имеет смысл отдельно рассчитать коэффициент прозрачности и предэкспоненциальный множитель. Имеем

$$\tau = \frac{a}{cD} \sqrt{\frac{2mc^2}{E}} = \frac{10^{-10}}{3 \cdot 10^8 \times D} \sqrt{\frac{2 \times 0,511 \cdot 10^6}{1,5}} \approx \frac{0,275 \cdot 10^{-15}}{D} \text{ с}.$$

Теперь осталось рассчитать коэффициент прозрачности:

$$D \approx \exp \left[ -\frac{4\pi d}{\lambda c} \sqrt{\frac{2(U - E)}{mc^2}} \right] = \exp \left[ -\frac{4\pi \times 3 \cdot 10^{-10}}{2,426 \cdot 10^{-12}} \sqrt{\frac{2 \times (2 - 1,5)}{0,511 \cdot 10^6}} \right] \approx 0,114.$$



Получаем окончательно

$$\tau \approx \frac{0,275 \cdot 10^{-15}}{0,114} \approx 2,4 \cdot 10^{-15} \text{ с} = 2,4 \text{ фс.}$$

Даже по масштабам микромира это время мало: прежде чем электрон просочится сквозь барьер, свет успеет пройти расстояние всего лишь 0,7 мкм. ■

Прозрачность барьера сильно зависит от энергии частицы в яме и от ширины и высоты барьера. Например, при увеличении ширины барьера в 2 раза новый коэффициент прозрачности будет равен, как легко догадаться, квадрату старого. Для электрона тогда получится значение  $D = 0,013$  и его время жизни в яме увеличится до  $\tau = 21$  фс. Это и объясняет отсутствие туннелирования в обычном мире с его высокими и широкими потенциальными барьерами.

**Задача 36.49.** Решить предыдущую задачу, если вместо электрона в ту же яму помещен протон.

**Решение.** Чтобы не решать аналогичную задачу с самого начала, можно воспользоваться результатами предыдущей. Протон массивнее электрона в  $\eta = 1836$  раз. В коэффициенте прозрачности масса частицы входит под квадратным корнем в показателе экспоненты. При изменении массы в  $\eta$  раз в показателе экспоненты появится множитель  $\sqrt{\eta}$  и новый коэффициент прозрачности будет равен старому, возведенному в степень  $\sqrt{\eta} = \sqrt{1836} \approx 42,85$ . Используя данные предыдущей задачи, получаем

$$D \approx (0,114)^{42,85} = 3,9 \cdot 10^{-41}.$$

Предэкспоненциальный множитель также умножится на  $\sqrt{\eta}$ , и время жизни протона в потенциальной яме будет равно:

$$\tau \approx \frac{42,85 \times 0,275 \cdot 10^{-15}}{D} \approx 3 \cdot 10^{26} \text{ с} \approx 10^{19} \text{ лет.}$$

Получилась невообразимо огромная величина, означающая, что протон будет жить в яме вечно: время существования Вселенной «всего» 13,7 млрд лет. ■

Эти две задачи демонстрируют сильную зависимость проницаемости барьера от массы частицы.

С эффектом туннелирования каждый из нас встречался в быту, хотя и не подозревал об этом. Если плотно скрутить две медные проволочки, то ток свободно проходит через место скрутки. А ведь медь покрыта на поверхности тонкой оксидной пленкой, являющейся изолятором. Как же тогда электроны проникают из одной проволочки в другую? Проводимость точек механического соединения проводников (скрутки проводов, зажимы, джамперы, электромагнитные реле и т. п.) обеспечивается туннелированием носителей зарядов.

Менее известен другой пример проявления туннельного эффекта, хотя от него зависит наша жизнь. Одним из источников энергии излучения Солнца и звезд является термоядерная цепная реакция синтеза ядер гелия из ядер водорода (протонов). Для этого протоны должны сблизиться на очень малые расстояния, чему препятствует кулоновское отталкивание между ними. Преодолеть отталкивание могут лишь протоны, обладающие очень высокими энергиями. Расчеты показывают, что температура плазмы в солнечном ядре должна составлять порядка 1 млрд градусов, чтобы энергия теплового движения могла сблизить протоны до этих расстояний и началась реакция синтеза. Но таких температур

на Солнце нет. К счастью, в результате туннельного эффекта протоны могут сблизиться на необходимое расстояние при гораздо более низких температурах — «всего лишь» около 10 млн градусов. О технических применениях туннельного эффекта в «рукотворных» устройствах речь пойдет в следующей части учебника.

Туннелирование через барьеры происходит с волнами любой природы. В рамках геометрической оптики из закона Снеллиуса вытекает, что при скользких углах падения преломленного луча не существует, т.е. свет не попадает из оптически более плотной среды в менее плотную (см. разд. 28.2). При более пристальном изучении этого вопроса в рамках теории электромагнетизма мы обнаружили в разделе 32.8, что преломленная волна все же существует, но экспоненциально затухает по мере удаления от поверхности раздела двух сред с глубиной проникновения порядка длины волны падающего света. Это явление, называемое *нарушенным полным внутренним отражением*, также имеет практические применения.

## 36.8 Оптическая аналогия прохождения частицы над барьером

В этом разделе мы покажем, что прохождение квантовой частицы через низкий потенциальный барьер аналогично отражению света на границе двух полубесконечных сред. Далее, прохождение частицы через потенциальный барьер конечной ширины может быть описано как множественное отражение классических волн, и в итоге опять-таки мы придем к известным результатам оптики. Целью данного раздела является демонстрация тесной связи различных областей физики.

### Ступенчатый потенциал

Этот потенциал изображен на рис. 36.2, а. Установление аналогии между квантовой механикой и светом означает, что мы хотим найти такие замены квантово-механических характеристик движения частицы на характеристики света, чтобы формулы квантовой механики перешли в соответствующие формулы для распространения света. Процедура замены будет изображена в формулах двойными стрелками, причем слева будут стоять квантово-механические величины, а справа — оптические. От этих формул следует отличать равенства, где по обе стороны стоят величины, относящиеся либо к частице, либо к световой волне.

Распространение квантовой частицы описывается в терминах ее волнового числа  $k = \sqrt{2mT}/\hbar = \sqrt{2m(E - U)}/\hbar$ , где  $T$  — кинетическая энергия частицы. Здесь и далее мы выписываем для частицы формулы, относящиеся к области барьера; соотношения для частицы вне барьера получаются при  $U = 0$ . Скорость частицы выражается соотношением

$$v = \frac{\hbar k}{m}. \quad (36.42)$$

Введем сначала показатель преломления  $n$  среды 2, соответствующей области барьера; его естественно определить как отношение скоростей частицы в областях 1 и 2:

$$\frac{\sqrt{E}}{\sqrt{E - U}} = \frac{1}{\sqrt{1 - U/E}} \longleftrightarrow n. \quad (36.43)$$

При  $U = 0$  получаем  $n = 1$  — показатель преломления вакуума.

В оптике волновое число связано с круговой частотой соотношением  $k = n\omega/c$ . При искомой замене волновое число частицы перейдет в волновое число света, т. е.

$$k = \frac{1}{\hbar} \sqrt{2m(E - U)} \longleftrightarrow \frac{n\omega}{c}. \quad (36.44)$$

Из соотношения  $\omega = kc/n$  для среды без дисперсии следует групповая скорость света  $d\omega/dk = c/n$ , в которую при искомой замене должна перейти скорость частицы  $v$ . Тогда уравнение (36.42) дает

$$\frac{\hbar k}{m} \longleftrightarrow \frac{c}{n}. \quad (36.45)$$

Разделив уравнение (36.44) на (36.45), найдем еще одну замену:

$$m \longleftrightarrow \frac{\hbar\omega}{c^2} n^2. \quad (36.46)$$

Конечно, «масса», стоящая здесь в левой части, ничего общего не имеет с массой фотона, которая равна нулю. Можно назвать ее «эффективной массой фотона в среде», и в вакууме при  $n = 1$  она равна  $\hbar\omega/c^2$ . Эта величина соответствует известной релятивистской связи масса—энергия, и она возникает при изучении влияния гравитационного поля на распространение световых лучей.

Указанные замены, как мы увидим, переводят формулы квантовой механики в формулы оптики.

Рассматривая прохождение частицы над низким потенциальным барьером, мы уже вывели коэффициент прохождения (36.36), который здесь обозначим через  $D_s$ :

$$D_s = \frac{4k_1 k_2}{(k_1 + k_2)^2} = \frac{4(k_1/k_2)}{(1 + k_1/k_2)^2},$$

$$k_1 = \frac{1}{\hbar} \sqrt{2mE}, \quad k_2 = \frac{1}{\hbar} \sqrt{2m(E - U)}. \quad (36.47)$$

Коэффициент отражения  $R_s = 1 - D_s$  (см. (36.35)) переписываем в виде

$$R_s = \left( \frac{k_1 - k_2}{k_1 + k_2} \right)^2 = \left( \frac{k_1/k_2 - 1}{k_1/k_2 + 1} \right)^2. \quad (36.48)$$

Применяя полученные выше замены, которые в данном случае сводятся к замене отношения  $k_1/k_2$  на показатель преломления  $n$ , мы получаем соответствующие коэффициенты прохождения и отражения для света, падающего перпендикулярно из вакуума на плоскую поверхность среды:

$$D_s = \frac{4n}{(n+1)^2}, \quad R_s = \left( \frac{n-1}{n+1} \right)^2. \quad (36.49)$$

Точно так же выглядят формулы Френеля (32.86) и (32.87), полученные в разделе 32.7 для относительных интенсивностей отраженного и преломленного света при угле падения  $90^\circ$ . Таким образом, прохождение частицы с энергией  $E$  над ступенчатым барьером высотой  $U$  эквивалентно нормальному падению света из вакуума на полубесконечную среду с показателем преломления

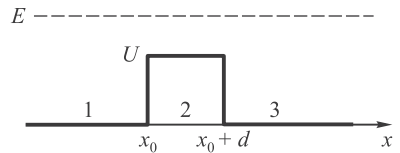
$$n = \frac{1}{\sqrt{1 - U/E}}. \quad (36.50)$$

Мы еще раз убедились, что физика (или природа) едина и квантовая механика имеет глубокие корни не только в классической механике, но и в волновой оптике.

## Прямоугольный барьер конечной ширины

Потенциальный барьер показан на рис. 36.8. Ранее мы рассматривали туннелирование сквозь такой барьер при  $E < U$ , теперь же изучим отражение и прохождение частицы с  $E > U$ . Решать задачу можно стандартным образом, записывая суперпозицию плоских волн для каждой из трех областей 1, 2 и 3 и «сшивая» затем решения, чтобы найти амплитуды волн. Однако мы заменим такой рутинный способ на классическое рассмотрение прохождения волн<sup>4</sup>, что проясняет физический смысл получаемого результата.

**Рис. 36.8.** Прохождение частицы над конечным прямоугольным барьером эквивалентно многократным отражениям от ступенчатого потенциала



Заметим прежде всего, что конечный барьер шириной  $d$  можно рассматривать как наложение двух ступенчатых барьеров, расположенных в точках  $x = x_0$  и  $x = x_0 + d$ . Это замечание дает возможность использовать выведенные ранее формулы.

Пусть квантовая волна с амплитудой, равной единице, движется слева направо и проникает в область над барьером в точке  $x_0$ . Вследствие частичного отражения ее амплитуда уменьшается и становится равной  $\sqrt{D_s}$ , где  $D_s$  — коэффициент прозрачности ступенчатого барьера. Далее она распространяется до точки  $x = x_0 + d$ , приобретая на пути сдвиг фазы  $\Delta\varphi = id(k_2 - k_1)$  по сравнению с фазой свободной частицы в этой же точке. Здесь волна снова встречается со ступенчатым барьером, в результате чего ее амплитуда вновь уменьшается в  $\sqrt{D_s}$  раз, и в итоге волна выходит за барьер с амплитудой

$$A^{(0)} = D_s \exp[id(k_2 - k_1)]. \quad (36.51)$$

Но мы учли только часть волны, выходящей наружу. Пришедшая в точку  $x = x_0 + d$  волна частично отражается от второго барьера (при этом ее амплитуда приобретает дополнительный множитель  $\sqrt{R_s}$ ), идет назад в точку  $x = x_0$ , снова отражается там (еще один множитель  $\sqrt{R_s}$  в амплитуде), возвращается назад в точку  $x = x_0 + d$ , где и выходит наружу. Полный пройденный этой частью волны путь равен  $3d$ , что дает сдвиг фазы  $\Delta\varphi = id(3k_2 - k_1)$ . В результате эта часть волны выходит за барьер с амплитудой

$$A^{(2)} = D_s R_s \exp[id(3k_2 - k_1)]. \quad (36.52)$$

Аналогичным образом происходят процессы с  $2n$  отражениями внутри барьера, и каждый из них приводит к волне с амплитудой

$$A^{(2n)} = D_s R_s^n \exp[i(2n + 1)dk_2 - idk_1]. \quad (36.53)$$

Амплитуда  $A_r$  результирующей волны получается суммированием выражения (36.53) по всем  $n$  от нуля до бесконечности:

$$A_r = \sum_{n \geq 0} A^{(2n)} = D_s e^{-id(k_1 - k_2)} \sum_{n \geq 0} \left( R_s e^{2idk_2} \right)^n = \frac{D_s e^{-id(k_1 - k_2)}}{1 - R_s e^{2idk_2}}. \quad (36.54)$$

<sup>4</sup>М. А. Smondyrev. On Optical Analogies for Quantum-mechanical Transmission Through Potential Barriers. J. Moscow Phys. Soc. **3**, 1–18 (1993).

Модуль амплитуды  $A_r$  прошедшей над барьером волны даст нам коэффициент прохождения  $D_r$ :

$$\begin{aligned} D_r &= \frac{D_s^2}{(1 - R_s e^{2idk_2})(1 - R_s e^{-2idk_2})} = \frac{(1 - R_s)^2}{1 + R_s^2 - 2R_s \cos 2dk_2} = \\ &= \frac{(1 - R_s)^2}{(1 - R_s)^2 + 4R_s \sin^2 dk_2}. \end{aligned} \quad (36.55)$$

Подставляя сюда квантово-механическое выражение (36.48) для  $R_s$ , получаем

$$D_r = \frac{4k_1^2 k_2^2}{4k_1^2 k_2^2 + (k_1^2 - k_2^2)^2 \sin^2(dk_2)}. \quad (36.56)$$

Стандартное решение уравнения Шрёдингера дает в точности такой же результат. Переходя к оптике, заменяем  $R_s$  на выражение (36.49) и  $k_2$  на  $\omega n/c$ . Получаем тогда коэффициент прозрачности пластины конечной толщины  $d$  при нормальном падении света с частотой  $\omega$ :

$$D_r = \left[ 1 + \left( \frac{n^2 - 1}{2n} \right)^2 \sin^2(d\omega n/c) \right]^{-1}. \quad (36.57)$$

Это выражение также в точности воспроизводит результат волновой оптики.

Подобным образом можно рассмотреть отраженную от барьера волну, но результат нам уже известен: коэффициент отражения от конечного барьера можно вычислить по формуле  $R_r = 1 - D_r$ .

Изучая формулы (36.56) и (36.57), мы обнаруживаем «окна прозрачности» при некоторых значениях частоты падающего света, когда  $D_r = 1$ ,  $R_s = 0$ , т.е. нет никакой отраженной волны. Это случается при  $dk_2 = \pi m$ ,  $m = 1, 2, 3, \dots$ , т.е. когда четное число полуволн (или целое число волн) света в среде укладывается в двойной ширине барьера:  $2dn = m\lambda$ . Это условие соответствует условию максимумов в проходящем свете при интерференции в тонких пленках. В обратном случае, когда двойная ширина барьера равна нечетному числу полуволн света в среде  $2dn = (m + 1/2)\lambda$ , мы приходим к минимальному значению коэффициента прозрачности:

$$(D_r)_{\min} = \frac{4k_1^2 k_2^2}{(k_1^2 + k_2^2)^2} = \frac{4n^2}{(n^2 + 1)^2}. \quad (36.58)$$

При повышении энергии частицы  $n \rightarrow 1$  (см. формулу (36.50)), так что как максимальные, так и минимальные значения  $D_r$  стремятся к единице.

В случае  $E < U$  мы имеем дело с туннелированием — частица «движется» внутри барьера с «мнимым» волновым числом:

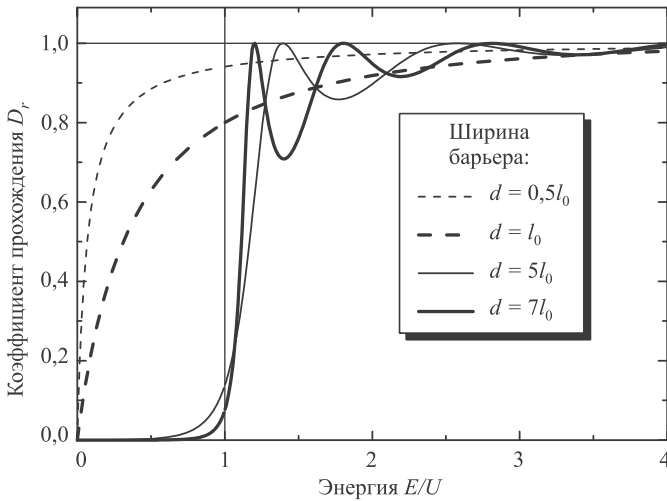
$$k_2 = i\kappa_2, \quad \kappa_2 = \frac{1}{\hbar} \sqrt{2m(U - E)}. \quad (36.59)$$

В этом случае тригонометрическая функция переходит в гиперболическую:

$$\frac{\sin(dk_2)}{k_2} = \frac{\text{sh}(d\kappa_2)}{\kappa_2},$$

и из уравнения (36.56) следует выражение для коэффициента прохождения

$$D_r = \frac{4k_1^2 \kappa_2^2}{4k_1^2 \kappa_2^2 + (k_1^2 + \kappa_2^2)^2 \text{sh}^2(d\kappa_2)}. \quad (36.60)$$



**Рис. 36.9.** Зависимость коэффициента прохождения частицы от относительной энергии  $E/U$  для прямоугольных барьеров разной ширины  $d$ , приведенной здесь в единицах параметра размерности длины  $l_0 = \hbar/\sqrt{2mU}$

Если, как это обычно бывает, аргумент  $d\kappa_2 \gg 1$ , то доминирует член с гиперболическим синусом, причем  $\text{sh}^2(d\kappa_2) \approx \exp(-2d\kappa_2)$ . Пренебрегая также предэкспоненциальными множителями, получаем уже знакомое выражение (36.38)

$$D_r \approx \exp\left(-\frac{2d}{\hbar}\sqrt{2m(U-E)}\right). \quad (36.61)$$

Результаты расчетов зависимости коэффициента прохождения от энергии частицы представлены на рис. 36.9 для потенциальных барьеров разной ширины, выраженной в единицах параметра размерности длины  $l_0 = \hbar/\sqrt{2mU}$ .

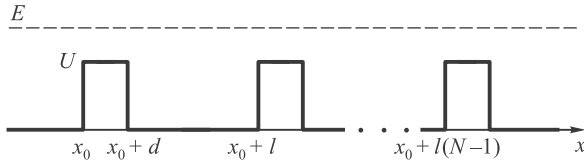
## Прохождение света через многослойную структуру

В этом разделе мы рассмотрим прохождение частицы над потенциальным барьером, составленным из  $N$  прямоугольных конечных барьеров шириной  $d$ , причем межбарьерное расстояние всюду одинаково и равно  $b$  (рис. 36.10).

В принципе решение задачи о нахождении коэффициента прохождения  $D^{(N)}$  через такой барьер может быть решена описанными выше способами. Сейчас нам важно только понять основной физический результат, так что нижеследующие формулы приводятся без вывода и только для справки. Выражение для коэффициента прохождения имеет вид

$$D^{(N)} = \left[1 + \left(\frac{1}{D_r} - 1\right) \frac{\sin[\text{h}](Nkl)}{\sin[\text{h}](kl)}\right]^{-1}. \quad (36.62)$$

Здесь  $D_r$  — коэффициент прохождения через одиночный прямоугольный барьер, вычисленный выше (при  $N = 1$  получаем из (36.62)  $D^{(1)} = D_r$ ). Величина  $k$  называется **квазиволновым числом** (для отличия от волновых чисел частицы  $k_1$  и  $k_2$  мы набираем ее прямым шрифтом).



**Рис. 36.10.** Многослойная структура с периодом  $l = d + b$ , образованная  $N$  прямоугольными потенциальными барьерами шириной  $d$  и межбарьерным расстоянием  $b$

В случае света мы используем для  $D_r$  результат (36.57). Связь квазиволнового числа с частотой падающего света и показателем преломления выражается тогда уравнением:

$$kl = \begin{cases} \arccos S, & \text{если } |S| \leq 1, \\ \ln(|S| + \sqrt{S^2 - 1}), & \text{если } |S| \geq 1, \end{cases} \quad (36.63)$$

где

$$S = \cos\left(d\frac{\omega}{c}n\right) \cos\left(b\frac{\omega}{c}\right) - \frac{n^2 + 1}{2n} \sin\left(d\frac{\omega}{c}n\right) \sin\left(b\frac{\omega}{c}\right). \quad (36.64)$$

Наконец, входящие в (36.62) функции определены как

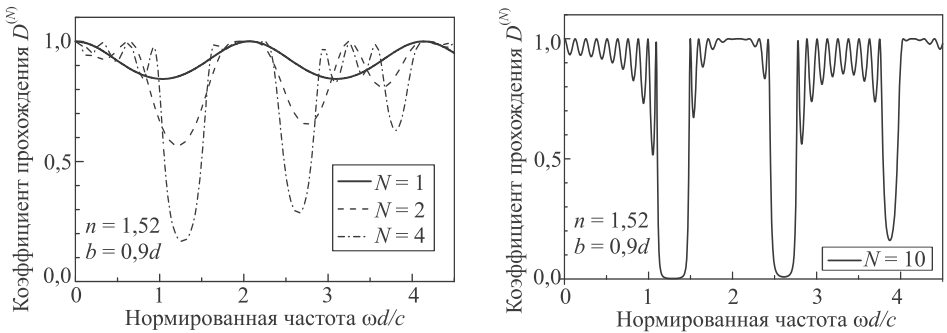
$$\sin[h]x = \begin{cases} \sin x, & \text{если } |S| \leq 1, \\ \text{sh } x, & \text{если } |S| \geq 1. \end{cases} \quad (36.65)$$

Ясно, что для частот, при которых абсолютно прозрачен прямоугольный барьер ( $D_r = 1$ ), будет прозрачен и наш составной барьер ( $D^{(N)} = 1$  в этом случае, как следует из формулы (36.62)). В дополнение появятся и новые окна прозрачности. Здесь ситуация очень похожа на случай с дифракционной решеткой (см. разд. 30.5). Численное решение показано на рис. 36.11, где для конкретности мы полагали  $n = 1,52$ ,  $b = 0,9d$ . Левый рисунок представляет собой коэффициент прохождения для одного, двух и четырех барьеров, правый — для десяти барьеров. Последний случай, когда число барьеров велико ( $N \gg 1$ ), представляет для нас особый интерес. Тенденции, обнаруживающиеся для  $N = 10$ , в пределе  $N \rightarrow \infty$  станут только отчетливее. А именно с ростом  $N$  некоторые минимумы становятся глубже и шире, и в пределе значение коэффициента прохождения в них стремится к нулю. Наоборот, амплитуда колебаний  $D^{(N)}$  при других частотах становится меньше, и коэффициент прохождения стремится к единице. Физическое объяснение этого явления заключается в том, что при выполнении определенных условий отраженные от концов прямоугольных барьеров волны взаимно компенсируют, гасят друг друга.

Тем самым мы пояснили физический смысл формирования при  $N \rightarrow \infty$  широких минимумов (так называемых *запрещенных зон*, непрозрачных для света или частицы) и максимумов с  $D \rightarrow 1$  (абсолютно прозрачных *разрешенных зон*).

Подчеркнем еще раз: для предельного случая периодической структуры с очень большим числом барьеров ( $N \rightarrow \infty$ ) зависимость коэффициента прохождения  $D^{(\infty)}$  от частоты падающего света такова, что:

- имеются целые полосы частот, в которых  $D^{(\infty)} = 0$ ; для таких частот составной барьер непрозрачен, свет этих частот полностью отражается от структуры;
- для других же частот, наоборот,  $D^{(N)} = 1$ , т. е. для них барьер полностью прозрачен, не происходит никакого отражения, свет свободно распространяется в такой структуре (на этом явлении основано изготовление просветленных объективов).



**Рис. 36.11.** Коэффициент прохождения в зависимости от частоты падающего света для малого  $N = 1, 2, 4$  (слева) и большого  $N = 10$  (справа) числа барьеров

Аналогичное явление происходит и для квантово-механической частицы, распространяющейся в периодическом потенциальном поле. При некоторых значениях энергии частицы бесконечная периодическая последовательность потенциальных барьеров становится для нее совершенно непрозрачной, даже если энергия частицы превышает высоту барьера. При других энергиях, наоборот, периодическая потенциальная структура становится для частицы абсолютно прозрачной. Так возникают запрещенные и разрешенные энергетические зоны в кристалле, и мы с ними в свое время познакомимся поближе.

## Контрольные вопросы

1. В чем заключается вероятностная интерпретация волновой функции?
2. Что такое условие нормировки волновой функции и в чем его физический смысл?
3. Запишите общее уравнение Шрёдингера. Какому соотношению классической физики оно соответствует?
4. Как действует на волновую функцию оператор импульса  $\hat{p}$ ?
5. Как действуют на волновую функцию оператор координаты  $\hat{r}$  и любая функция этого оператора  $f(\hat{r})$ ?
6. Что такое гамильтониан системы?
7. Запишите стационарное уравнение Шрёдингера.
8. При решении стационарного уравнения Шрёдингера для некоторой системы получается ряд значений энергии  $E_1, E_2, E_3, \dots$ . Может ли экспериментатор получить какое-то иное значение при измерении энергии этой системы?
9. Дайте наглядное объяснение, почему квантуется энергия частицы, находящейся в потенциальной яме с бесконечно высокими стенками.
10. Почему в выражении (36.26) для уровней энергии в потенциальной яме с бесконечно высокими стенками отсутствуют отрицательные значения квантового числа  $n$ ?
11. Что такое вырождение энергетических уровней?
12. Что такое нулевые колебания квантового осциллятора? Почему их существование неизбежно?



13. Сформулируйте принцип соответствия Бора.
14. В бесконечно глубокой потенциальной яме расстояние между уровнями возрастает с ростом номера уровня  $n$  (см. (36.26)). В то же время, казалось бы, принцип соответствия Бора требует, чтобы уровни сгущались к классическому значению  $h\nu_{cl}$ . Покажите, что принцип соответствия справедлив и для этой системы.
15. Пусть частица налетает на потенциальный барьер, причем ее энергия больше высоты барьера:  $E > U$ . Можно ли по результату такого процесса обнаружить разницу между предсказаниями классической и квантовой теорий?
16. Что такое эффект туннелирования? Почему он не наблюдается при движении макроскопических тел в нашей повседневной жизни?
17. Металлическая шайба массой  $m = 10$  г скользит по гладкой поверхности со скоростью  $v_0 = 1$  м/с. На ее пути стоит небольшая горка с гладкими склонами. Высота горки  $H = 5,5$  см, ее ширина  $d = 10$  см. Как известно из курса механики, начальная скорость шайбы позволяет ей подняться лишь на высоту  $H_0 = v^2/2g = 1/(2 \times 9,8) = 5,1 \cdot 10^{-2}$  м = 5,1 см. Квантовая же механика утверждает, что в принципе шайба может оказаться по другую сторону горки. Оцените вероятность такого события.
18. Как объяснить, что при туннелировании частица может попасть в область, где ее полная энергия меньше потенциальной?
19. Квантовая механика дает только вероятностное описание движения частиц, говорить об их положении, скорости и т.п. смысла не имеет. Почему тогда эта теория может широко использоваться при проектировании, например, устройств, принцип действия которых основан на квантовых свойствах систем?

## Глава 37

# Теория атома

Раз уж мы получили в руки такое мощное оружие, как уравнение Шрёдингера, то естественно вернуться к атому, начав с простейшего случая — атома водорода. Надо же убедиться, что квантовая механика приведет к тем же результатам, что и полуклассическая теория атома по Бору. К тому же есть надежда, что новая теория подарит нам неожиданные открытия. Жаль только, что изучение способов решения уравнения Шрёдингера не входит в рамки настоящего курса. Но не беда: мы постараемся обойтись без излишней математики, угадывая свойства решений на основе интуиции, выработанной при изучении классической физики. Следует только помнить, что читателя никто обманывать не собирается: все «угаданное» может быть получено из уравнения Шрёдингера точными математическими расчетами.

### 37.1 Коммутирующие операторы

В предыдущей главе мы установили, что классические динамические переменные заменяются в квантовой механике операторами, действующими на волновую функцию. Результаты измерения некоей величины  $A$  всегда будут определяться собственными значениями соответствующего оператора:  $\hat{A}\psi_n = A_n\psi_n$ . Если система находится в каком-нибудь собственном состоянии  $\psi_n$  оператора  $A$ , то измерение наверняка дает собственное значение  $A_n$ . Если же система находится в каком-то другом состоянии, то измерение  $A$  с определенной вероятностью дает какое-то из собственных значений, причем эта вероятность зависит от волновой функции состояния и, разумеется, от измеряемой величины  $A$ . Сказанное регулируется следующим правилом.

**Правило 3.** ♦ *Если система находится в состоянии, волновая функция которого является суперпозицией собственных функций  $\psi_n$  какого-то оператора  $A$ , т. е. представляется в виде*

$$\psi = \sum_n c_n \psi_n, \quad \sum_n |c_n|^2 = 1, \quad (37.1)$$

*то измерение величины  $A$  с вероятностью  $|c_n|^2$  даст значение  $A_n$ .* ♦

Операторы в квантовой механике таковы, что любую функцию можно разложить в ряд вида (37.1).

Пусть система находится в состоянии с определенным значением величины  $A$ . Это значит, что ее волновая функция является собственной функцией оператора  $\hat{A}$ . Может ли другая величина  $B$  также иметь определенное значение, т. е. может ли состояние быть собственным сразу и для оператора  $\hat{A}$ , и для оператора  $\hat{B}$ ?

◆ *Два оператора  $\hat{A}$  и  $\hat{B}$  имеют общий набор собственных состояний тогда и только тогда, когда они коммутируют:*

$$\boxed{\hat{A}\hat{B} = \hat{B}\hat{A}.} \quad \blacklozenge$$

Пусть, например, волновая функция  $\psi_n$  является собственной функцией оператора  $\hat{A}$  с собственным значением  $A_n$  и собственной функцией оператора  $\hat{B}$  с собственным значением  $B_n$ . Подействуем на нее сначала оператором  $\hat{A}$ , а затем оператором  $\hat{B}$ :

$$\hat{B}\hat{A}\psi_n = A_n\hat{B}\psi_n = A_nB_n\psi_n. \quad (37.2)$$

Поменяем теперь порядок действия операторов:

$$\hat{A}\hat{B}\psi_n = B_n\hat{A}\psi_n = B_nA_n\psi_n. \quad (37.3)$$

Величины  $A_n, B_n$  — обычные числа, от перестановки которых произведение не меняется. Стало быть, совпадают правые части равенств (37.2) и (37.3). Тогда совпадают и их левые части, т. е.  $\hat{A}\hat{B}\psi_n = \hat{B}\hat{A}\psi_n$ . Так как любую функцию можно представить в виде суперпозиции (37.1), то отсюда следует, что результат последовательного действия операторов  $\hat{A}$  и  $\hat{B}$  на любую функцию  $\psi$  не зависит от их порядка.

Можно сказать иначе: если результат последовательного действия двух операторов не зависит от порядка их действия, то величины, определяемые этими операторами, могут одновременно иметь определенные значения. Но так бывает далеко не всегда. Рассмотрим пример:

$$\begin{aligned} \hat{p}_x\hat{x}\Psi &= -i\hbar\frac{\partial}{\partial x}x\Psi = -i\hbar(\Psi + x\Psi'), \\ \hat{x}\hat{p}_x\Psi &= -i\hbar x\frac{\partial}{\partial x}\Psi = -i\hbar x\Psi', \end{aligned}$$

т. е. для любой функции  $\Psi$

$$(\hat{p}_x\hat{x} - \hat{x}\hat{p}_x)\Psi = -i\hbar\Psi,$$

или просто

$$\boxed{\hat{p}_x\hat{x} - \hat{x}\hat{p}_x = -i\hbar.}$$

*Поскольку эти операторы не коммутируют, то координата и проекция импульса на ту же ось не измеримы одновременно.* Этот вывод и есть истинный источник соотношения неопределенностей Гейзенберга, физический смысл которых разобран выше.

Особое значение имеет свойство коммутации операторов с гамильтонианом, т. е. с оператором полной энергии  $\hat{H}$ . Если какой-то оператор  $\hat{A}$  коммутирует с  $\hat{H}$ , то существует общее собственное состояние. Будучи собственным состоянием гамильтониана, оно стационарно по определению. В стационарном же состоянии система пребывает неограниченно долго. Это означает одновременное сохранение и энергии, и величины  $A$ . Таким образом, утверждение о сохранении некой величины эквивалентно тому, что она может иметь определенное значение вместе с энергией, т. е. соответствующий ей оператор коммутирует с гамильтонианом.

## 37.2 Момент количества движения

В классической механике момент количества движения (его также называют моментом импульса или угловым моментом) выражается в виде векторного произведения радиуса-вектора на импульс:  $\vec{L} = [\vec{r} \times \vec{p}]$ . Такая же связь соблюдается между операторами квантовой механики:

$$\hat{L} = [\hat{r} \times \hat{p}] = -i\hbar[\vec{r} \times \vec{\nabla}].$$

Это приводит к выражению для  $z$ -компонента момента импульса

$$\hat{L}_z = -i\hbar \left( x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \quad (37.4)$$

и аналогичным выражениям для других компонентов. В разделе 36.3 обсуждалось, почему оператор проекции момента импульса на какую-то ось связан с производной по углу поворота вокруг этой оси (см. уравнение (36.15)). В сферических координатах поворот вокруг оси  $z$  эквивалентен сдвигу по азимутальному углу  $\varphi$ , и потому оператор (37.4) имеет особенно простой вид

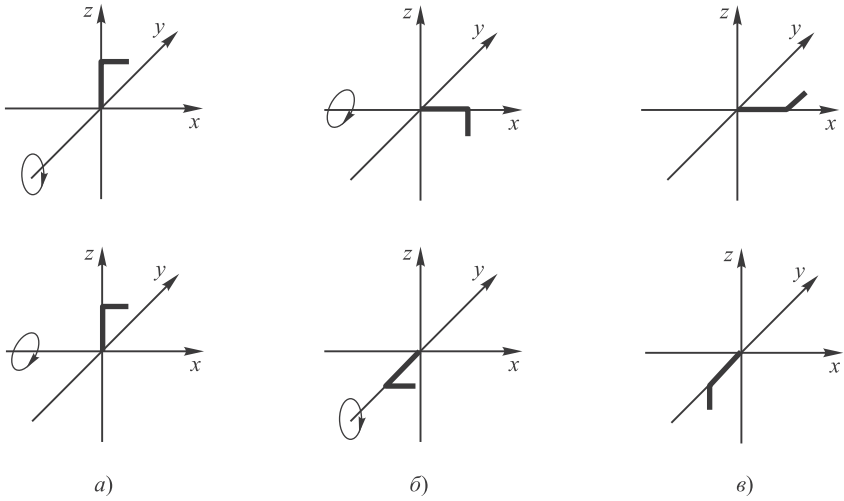
$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \varphi}. \quad (37.5)$$

Выражения для других компонентов  $\hat{L}_x$  и  $\hat{L}_y$  в сферических координатах довольно сложны, они зависят как от полярного угла  $\theta$ , так и от азимутального угла  $\varphi$ . Мы выпишем здесь лишь оператор квадрата момента количества движения:

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 = -\hbar^2 \left( \frac{\partial^2}{\partial \theta^2} + \text{ctg } \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right). \quad (37.6)$$

Выражение (37.6) также достаточно сложно, и мы его практически использовать не будем. Но даже только глядя на него, уже можно сделать важные выводы.

- В оператор  $\hat{L}^2$  входит не сам угол  $\varphi$ , а лишь производная по нему. Это означает, что  $\hat{L}^2$  коммутирует с оператором  $\hat{L}_z$ .
- Так как ось  $z$  ничем не выделена, оператор квадрата момента импульса коммутирует и с операторами проекции момента импульса на любую другую ось (в частности, с  $\hat{L}_x$  и  $\hat{L}_y$ ).



**Рис. 37.1.** В верхнем ряду Г-образная фигура сначала поворачивается по часовой стрелке на  $90^\circ$  вокруг оси  $y$  (а), затем поворачивается (также по часовой стрелке на  $90^\circ$ ) вокруг оси  $x$  (б); в нижнем ряду выбрана обратная последовательность тех же поворотов. Конечные результаты (в) различаются, что демонстрирует некоммутативность операторов поворота

- Из невыписанных здесь выражений для  $\hat{L}_x$  и  $\hat{L}_y$  следует, что операторы  $\hat{L}_x, \hat{L}_y, \hat{L}_z$  не коммутируют между собой.

Вместо формального математического доказательства последнего утверждения укажем источник этого свойства. Напомним, что  $\hat{L}_x, \hat{L}_y, \hat{L}_z$  являются операторами поворота системы вокруг осей  $x, y, z$  соответственно. Но результат двух таких поворотов зависит от их последовательности (рис. 37.1).

- Из сказанного вытекает важное следствие: одновременно измеримы лишь квадрат момента импульса и одна из его проекций (в качестве таковой обычно выбирают  $\hat{L}_z$ ). Это значит, что вектор  $\vec{L}$  в квантовой механике не имеет определенного направления и его нельзя считать классическим вектором с тремя определенными компонентами.

Таким образом, «квантовый момент количества движения» можно условно представить себе как вектор с фиксированной длиной (определенное значение квадрата момента импульса), направленный под фиксированным углом к оси  $z$  (определенное значение проекции), но прецессирующий вокруг этой оси, как показано на рис. 37.2 справа. При этом другие компоненты импульса не определены. Это не более чем механическая аналогия, но она верно отражает существенные свойства момента импульса в квантовой механике.

Найдем теперь собственные функции и значения оператора  $\hat{L}_z$ . Из уравнения

$$-i\hbar \frac{\partial}{\partial \varphi} \Phi(\varphi) = L_z \Phi(\varphi),$$

откуда

$$\Phi(\varphi) = Ae^{iL_z\varphi/\hbar}.$$

При повороте на угол  $2\pi$  система возвращается в первоначальное состояние. Чтобы волновая функция  $\Phi(\varphi)$  не изменилась, необходимо выполнение условия  $L_z = \hbar m$ , где  $m$  — целое (необязательно положительное) число. Константа  $A$  определяется условием нормировки: интеграл от  $|\Phi(\varphi)|^2 = |A|^2$  по углу  $\varphi$ , изменяющемуся от 0 до  $2\pi$ , должен быть равен единице, откуда  $A = 1/\sqrt{2\pi}$ .

Таким образом, мы приходим к квантованию: проекция момента импульса  $L_z$  может принимать лишь целые значения в единицах постоянной Планка:

$$\boxed{L_z = \hbar m, \quad (m = 0, \pm 1, \pm 2, \dots).} \quad (37.7)$$

Число  $m$  называют **магнитным квантовым числом** (смысл названия станет понятен позже). Собственная волновая функция оператора  $\hat{L}_z$ , соответствующая данному значению  $m$ , имеет вид

$$\Phi_m(\varphi) = \frac{1}{\sqrt{2\pi}} e^{im\varphi}. \quad (37.8)$$

По сути, волновая функция (37.8) описывает стоячую волну вдоль окружности. Роль координаты играет угол  $\varphi$ , роль волнового числа — **магнитное квантовое число  $m$** . Но значения переменной  $\varphi$  ограничены пределами 0 и  $2\pi$ . Наша «круговая» волна как будто заключена в потенциальную яму и совершает финитное движение. Отсюда — квантование проекции момента импульса в соответствии с установленными выше законами квантовой механики.

Найдем теперь правила квантования квадрата момента импульса. Решение соответствующего уравнения на собственные функции оператора  $\hat{L}^2$  достаточно сложно, и мы заменим его не очень строгими, но простыми соображениями. Пусть в какой-то системе максимальная величина магнитного квантового числа  $m$  равна целому неотрицательному числу  $l$ . Тогда минимальное значение  $m$ , очевидно, равно  $-l$ , так что  $m$  пробегает  $2l + 1$  возможных значений:

$$\boxed{m = -l, -l + 1, \dots, -1, 0, 1, \dots, l - 1, l.}$$

В классическом случае максимально возможная проекция момента импульса совпадает с модулем вектора  $\vec{L}$ . Но не следует ожидать, что оператор  $\hat{L}^2$  будет иметь собственные значения  $\hbar^2 l^2$ . Мы уже знаем, что даже при максимальной величине проекции момент импульса не параллелен оси  $z$  (иначе были бы известны все три компоненты момента). Стало быть, собственные значения оператора  $\hat{L}^2$  должны быть больше  $\hbar^2 l^2$ . Чему же они равны?

Если в пространстве нет выделенного направления, то любое значение  $m$  равновероятно и среднее значение квадрата проекции момента на ось  $z$  равно

$$\langle \hat{L}_z^2 \rangle = \frac{1}{2l + 1} \sum_{m=-l}^l \hbar^2 m^2 = \frac{2\hbar^2}{2l + 1} \sum_{m=1}^l m^2 = \frac{2\hbar^2}{2l + 1} \frac{l(l + 1)(2l + 1)}{6} = \hbar^2 \frac{l(l + 1)}{3}.$$

При выводе использовалась известная формула для суммы квадратов целых чисел. Заметим, что все три координатные оси равноправны, следовательно, тот же результат справедлив для средних значений квадратов операторов всех проекций момента импульса, т. е.

$$\langle \hat{L}_x^2 \rangle = \langle \hat{L}_y^2 \rangle = \langle \hat{L}_z^2 \rangle = \hbar^2 \frac{l(l+1)}{3}.$$

Поэтому их сумма дает квадрат оператора момента импульса, среднее значение которого равно, таким образом,

$$\boxed{\langle \hat{L}^2 \rangle = \hbar^2 l(l+1)}. \quad (37.9)$$

Именно такой формулой описываются собственные значения оператора квадрата момента импульса, так что условно можно считать, что длина вектора  $\vec{L}$  в квантовой механике равна  $\hbar\sqrt{l(l+1)}$ . Целое неотрицательное квантовое число  $l$  называют **азимутальным квантовым числом**.

Для сравнения получим тем же способом классический ответ. Если  $l$  — максимальное значение  $m$  для классического вектора, то  $m$  пробегает непрерывный ряд значений от  $-l$  до  $l$  с равной вероятностью  $dm/(2l)$ . Разница в том, что из-за непрерывности сумма заменяется интегралом, и мы получаем для средних квадратов всех трех проекций момента импульса выражения

$$\langle L_x^2 \rangle + \langle L_y^2 \rangle + \langle L_z^2 \rangle = \int_{-l}^l \frac{dm}{2l} (\hbar^2 m^2) = \frac{\hbar^2}{2l} \int_{-l}^l dm m^2 = \hbar^2 \frac{l^2}{3}.$$

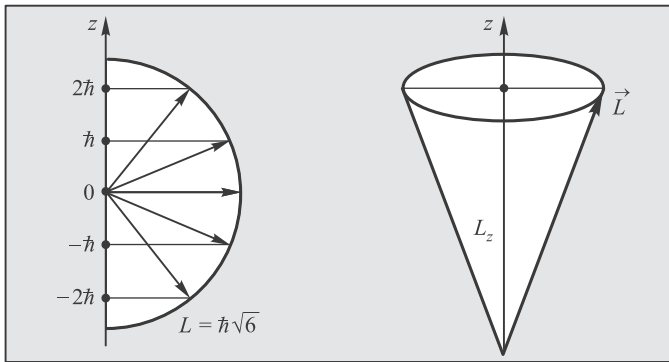
Складывая их, приходим к обычному результату классической физики:

$$\langle \vec{L}^2 \rangle = \hbar^2 l^2.$$

При больших значениях  $l$  оба результата совпадают (принцип соответствия Бора).

Главный итог этого раздела — знакомство с правилами квантования момента импульса: собственное значение квадрата момента импульса определяется величиной азимутального квантового числа  $l$ , а проекция момента импульса — величиной магнитного квантового числа  $m$ , которое может принимать любое из значений  $0, \pm 1, \dots, \pm l$ . Если все-таки пытаться представить себе квантовый вектор момента количества движения как обычный вектор, то можно сказать, что при данной длине  $L$  этого вектора он составляет с выделенной осью лишь строго определенные углы (рис. 37.2), так как его проекция  $L_z$  принимает строго определенный ряд значений. Отсутствие определенных значений у других проекций —  $L_x$  и  $L_y$  — можно представить себе как результат непрерывной прецессии вектора углового момента вокруг оси  $z$ , подобной прецессии гироскопа (детской юлы, например). Подчеркнем еще раз, что эта картинка — всего лишь попытка изобразить квантовые свойства в классических образах.

**Задача 37.50.** Показать, что согласно квантовой механике направление момента импульса  $\vec{L}$  не может совпадать с выделенным в пространстве направлением и что в пределе больших азимутальных чисел  $l \gg 1$  восстанавливаются классические свойства.



**Рис. 37.2.** Слева показаны возможные положения вектора момента импульса при  $l=2$ : длина вектора равна  $\hbar\sqrt{6}$ , а его проекция на выделенную ось может принимать только значения  $0, \pm\hbar, \pm 2\hbar$ ; справа — классический образ вектора с фиксированными длиной  $L$  и проекцией  $L_z$ , прецессирующего вокруг оси  $z$ , так что проекции  $L_x, L_y$  не имеют определенных значений

**Решение.** Поскольку модуль вектора момента импульса принимает значения  $\hbar\sqrt{l(l+1)}$ , а его проекция на выделенное направление равна  $\hbar m$ ,  $m = -l, -l+1, \dots, l-1, l$ , можно ввести угол  $\theta$  между направлением момента импульса и выделенной осью, так что  $\cos \theta$  будет принимать лишь определенные значения:

$$\cos \theta = \frac{m}{\sqrt{l(l+1)}}.$$

Отсюда следует, что минимальное значение угла  $\theta$  определяется максимальным значением его косинуса, достигаемым при  $m = l$ :

$$\cos \theta_{\min} = \sqrt{\frac{l}{l+1}}.$$

Видно, что при любом конечном значении  $l$  угол не равен нулю. Например, для состояний с  $l = 1$  имеем  $\cos \theta_{\min} = 1/\sqrt{2}$ , т. е.  $\theta_{\min} = 45^\circ$ , а для состояний с  $l = 3$  находим  $\cos \theta_{\min} = \sqrt{3}/2$  и  $\theta_{\min} = 30^\circ$ . Видно, что с ростом  $l$  минимальный угол между моментом импульса и осью уменьшается, и в пределе  $l \rightarrow \infty$  получаем  $\theta_{\min} = 0$ . Это и есть классическое свойство момента импульса: способность быть в точности параллельным любому выделенному направлению. ■

### 37.3 Атом водорода

Запишем стационарное уравнение Шрёдингера для атома водорода:

$$-\frac{\hbar^2}{2m_e} \Delta \psi(\vec{r}) - \frac{e^2}{4\pi\epsilon_0 r} \psi(\vec{r}) = E\psi(\vec{r}). \quad (37.10)$$

Для данной задачи оператор Лапласа удобно взять в сферических координатах  $r, \theta, \phi$ , и тогда уравнение Шрёдингера (37.10) примет вид

$$\left[ -\frac{\hbar^2}{2m_e} \left( \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{\hat{L}^2}{2m_e r^2} - \frac{e^2}{4\pi\epsilon_0 r} \right] \psi(r, \theta, \varphi) = E\psi(r, \theta, \varphi). \quad (37.11)$$

Мы не станем решать это уравнение, но просто внимательно на него посмотрим.



Заметим, что та часть оператора Лапласа в (37.11), которая зависит от углов  $\theta, \phi$  и производных по ним, входит только в состав оператора квадрата момента импульса (37.6). Довольно ясен физический смысл этого члена. Представим себе, что в поле центральных сил по орбите радиусом  $r$  движется классическая частица с импульсом  $\vec{p}$ . Ее момент количества движения равен  $L = |[\vec{r} \times \vec{p}]| = rp_{\perp}$ , где  $p_{\perp} = m_e v_{\perp}$  — проекция импульса на направление, ортогональное радиусу-вектору  $\vec{r}$ . Обозначим через  $T_{\perp} = p_{\perp}^2 / (2m_e)$  кинетическую энергию «ортогонального» движения. Ее можно выразить через квадрат момента количества движения:

$$T_{\perp} = \frac{L^2}{2m_e r^2}.$$

Этот член добавляется к потенциальной энергии кулоновского притяжения к ядру, и его можно интерпретировать как потенциальную энергию в поле центробежных сил. Действительно, если  $T_{\perp}$  — потенциальная энергия, то ее производная по  $r$  должна дать соответствующую силу:

$$F = -\frac{d}{dr} \left( \frac{L^2}{2m_e r^2} \right) = \frac{L^2}{m_e r^3} = \frac{p_{\perp}^2}{m_e r} = \frac{m_e v_{\perp}^2}{r}.$$

В конечном выражении легко узнать известную из классической механики формулу для центробежной силы. Квантовая механика, как это и должно быть, воспроизводит на новом уровне результаты классической: теперь момент импульса стал оператором, но вошел на прежних правах в выражение для оператора полной энергии (гамильтониана).

Любой оператор коммутирует сам с собой, и так как оператор квадрата углового момента (37.6) вообще не зависит от радиальной переменной  $r$ , то  $\hat{L}^2$  коммутирует с гамильтонианом уравнения Шрёдингера (37.11). Кроме того, оператор проекции момента импульса  $\hat{L}_z$  коммутирует с  $\hat{L}^2$  и, стало быть, с гамильтонианом. Следовательно, выполняются классические законы сохранения квадрата и одной проекции момента импульса. Эти законы сохранения справедливы для любого центрально-симметричного поля: специфика кулоновского взаимодействия пока нами не использовалась. Поэтому проекция и квадрат момента могут быть определены одновременно с энергией, и волновая функция стационарного состояния будет зависеть от квантовых чисел  $l$  и  $m$ . Однако в уравнении Шрёдингера (37.11) гамильтониан вовсе не зависит от оператора проекции момента импульса. Это значит, что энергия состояния не будет зависеть от магнитного квантового числа  $m$ . Иными словами, в любом центрально-симметричном поле имеется вырождение по  $m$ , кратность которого равна  $2l + 1$ .

Мы уже знаем, что источником вырождения должна служить та или иная симметрия. В классической физике движение частицы в центрально-симметричном поле всегда происходит по орбите, лежащей в одной плоскости. Но сама эта плоскость может быть произвольной в зависимости от начального положения и скорости частицы. Ясно, что значение полной энергии частицы не зависит при этом от ориентации плоскости орбиты в пространстве. Это и есть искомая симметрия, приводящая к вырождению по магнитному квантовому числу.

В кулоновском поле (равно как и в гравитационном) имеется еще одно специфическое вырождение, приводящее к тому, что энергия системы не зависит и от

квантового числа  $l$ . Вспомним опять классическую физику. В кулоновском поле финитное движение частицы совершается только по эллипсу. Возьмем в качестве аналогии искусственный спутник. Поместим его на каком-то расстоянии от Земли (т. е. зададим потенциальную энергию) и сообщим ему какую-то скорость (зададим кинетическую энергию). Таким образом, мы задали полную энергию спутника. Но определена ли его орбита? Разумеется, нет! При той же полной энергии направление начальной скорости влияет на форму орбиты — от прямой линии (вертикальное падение) при нулевом моменте импульса до максимально возможного при данной полной энергии. Нулевой момент соответствует чисто радиальным колебаниям сквозь центр притяжения, когда вовсе нет кругового движения и эллипс вырождается в прямую линию (для спутника такое колебание невозможно, но микрочастицы — иное дело). Максимально возможный момент импульса достигается в обратном случае чисто круговой орбиты, когда совсем нет радиального движения. Важно, что величина этого максимального момента зависит от полной энергии спутника.

Итак, классическая физика подсказывает нам следующие свойства решений уравнения Шрёдингера:

- энергия уровней не должна зависеть (вырождение) от квантовых чисел  $l, m$ ;
- квантовое число  $l$  должно пробегать ряд целых значений от нулевого до максимального;
- максимальное значение  $l$  должно зависеть от энергии электрона.

Вооружившись знанием классической механики, мы можем смело приступить к изучению квантовой. Теперь станут понятны свойства решений уравнения Шрёдингера для атома водорода. Его решениями являются волновые функции, нумеруемые тремя квантовыми числами  $n, l, m$ . Про  $l$  и  $m$  уже много говорилось, а  $n$  — знакомое нам по атому Бора главное квантовое число, принимающее целые положительные значения. Разным наборам чисел  $n, l, m$  отвечают разные волновые функции, вид которых мы обсудим позже.

**Задача 37.51.** Доказать, что волновая функция  $\psi(r) = C \exp(-r/a)$  удовлетворяет уравнению Шрёдингера для атома водорода. Найти энергию этого состояния, параметр  $a$  и нормировочную постоянную  $C$ .

**Решение.** Вычисляем производные волновой функции по  $r$ :

$$\frac{d\psi(r)}{dr} = -\frac{1}{a}\psi(r), \quad \frac{d^2\psi(r)}{dr^2} = \frac{1}{a^2}\psi(r)$$

и подставляем их в (37.11). Поскольку функция  $\psi(r)$  не зависит от углов  $\theta, \phi$ , действие на нее дифференциального оператора квадрата углового момента дает ноль. Приходим тогда к алгебраическому уравнению

$$E = -\frac{\hbar^2}{2m_e a^2} + \frac{\hbar^2}{m_e a r} - \frac{e^2}{4\pi\epsilon_0 r}. \quad (37.12)$$

Левая часть уравнения (37.12) не зависит от радиальной переменной  $r$ . Следовательно, в правой части уравнения должны взаимно уничтожиться слагаемые, зависящие от  $r$ , откуда находим параметр  $a$ :

$$a = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2}. \quad (37.13)$$

Мы получили выражение (35.4) для радиуса Бора:  $a = a_B$ . Теперь из (37.12) следует выражение для энергии этого состояния

$$E = -\frac{\hbar^2}{2m_e a_B^2} = -\frac{m_e e^4}{2(4\pi\epsilon_0)^2 \hbar^2}, \quad (37.14)$$

которое совпадает с энергией основного состояния ( $n = 1$ ) по модели атома Бора (см. формулу (35.5)).

Найдем нормировочную постоянную. Интеграл от квадрата модуля волновой функции по всему пространству должен быть равен единице. Интеграл вычисляется в сферических координатах, в которых элемент объема равен  $dV = r^2 \sin\theta dr d\theta d\varphi$ . Так как волновая функция  $\psi(r)$  не зависит от  $\theta$  и  $\varphi$ , интегрирование по углам дает  $4\pi$ , и можно записать цепочку равенств:

$$\int |\psi(r)|^2 d\vec{r} = |C|^2 \int_0^\infty r^2 e^{-2r/a_B} dr \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\varphi = 4\pi |C|^2 \int_0^\infty r^2 e^{-2r/a_B} dr = |C|^2 \pi a_B^3 = 1. \quad (37.15)$$

Отсюда находим:  $C = 1/\sqrt{\pi a_B^3}$  — и получаем окончательное выражение для волновой функции основного состояния атома водорода:

$$\boxed{\psi(r) = \frac{1}{\sqrt{\pi a_B^3}} e^{-r/a_B}} \blacksquare \quad (37.16)$$

У найденной нами волновой функции нет узлов, так что это действительно волновая функция основного состояния (помимо того что она соответствует энергии нижнего уровня в модели атома Бора с главным квантовым числом  $n = 1$ ). Отсутствие зависимости от углов  $\theta, \phi$  означает, что эта волновая функция описывает состояния с  $l = 0$  и  $m = 0$ . Поскольку основное состояние всегда невырождено, состояний с  $n = 1$  и другими значениями азимутального квантового числа  $l$ , кроме  $l = 0$ , не существует. Выше, рассматривая классическую частицу в кулоновском поле, мы угадали сам факт существования максимального значения  $l_{\max}$  азимутального квантового числа. Теперь мы заметили, что оно связано с главным квантовым числом соотношением  $l_{\max} = n - 1$ , которое остается справедливым и для других решений уравнения Шрёдингера, соответствующих возбужденным уровням энергии атома водорода.

Для энергетических уровней из уравнения Шрёдингера получается в точности такая же формула, что и в модели атома Бора:

$$E_n = -\frac{m_e e^4 Z^2}{2(4\pi\epsilon_0)^2 \hbar^2 n^2}. \quad (37.17)$$

Как видно, энергия действительно не зависит от квантовых чисел  $l, m$ . При этом, как следует из свойств решений уравнения (37.11), **азимутальное квантовое число  $l$  принимает целые значения от 0 до  $n - 1$** . Таким образом, существует максимальное значение орбитального момента, и оно зависит от энергии электрона. И это свойство, угаданное нами на основе классической физики, воспроизвелось в квантовой механике.

Удивительно, как квантовая механика, низвергнувшая столько классических представлений, дает аналогичные классическим результаты там, где в дело

вступают свойства симметрии системы. Отсюда вывод: симметрия играет более важную роль, чем конкретные физические законы. Когда-нибудь будут открыты новые законы, которые обобщат и квантовую механику, и все теории, которые ныне находятся на переднем крае науки. Но свойства симметрии системы так или иначе проявят себя и тогда.

Квадрат модуля волновой функции  $|\psi_{nlm}(\vec{r})|^2$  дает нам плотность вероятности найти электрон в бесконечно малом объеме  $dV$ , окружающем точку  $\vec{r}$ . Для состояний с  $l = 0, m = 0$  волновая функция сферически симметрична, поэтому можно найти вероятность  $dW(r)$  того, что электрон обнаружится в шаровом слое радиусом  $r$  и толщиной  $dr$ , просто умножив плотность вероятности на объем шарового слоя:

$$dW(r) = 4\pi r^2 |\psi_{n00}(r)|^2 dr. \quad (37.18)$$

Величина  $dW(r)/dr$  представляет собой плотность вероятности найти электрон на расстоянии  $r$  от ядра. Наиболее вероятным называется расстояние, соответствующее максимуму плотности вероятности.

**Задача 37.52.** Найти наиболее вероятное расстояние  $r_1$  электрона от ядра в основном состоянии атома водорода.

**Решение.** Нам надо найти максимум функции  $r^2 \exp(-2r/a_B)$ . Дифференцируем это выражение по  $r$  и приравняем производную нулю. Получаем

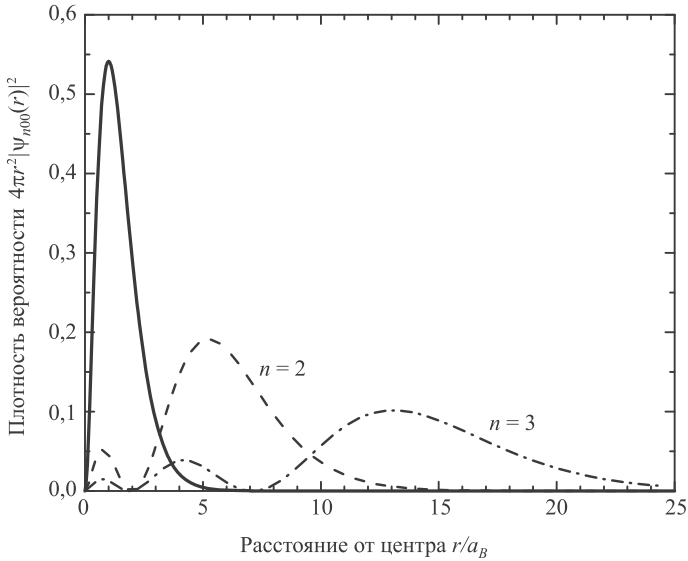
$$2re^{-2r/a_B} - r^2 \frac{2}{a_B} e^{-2r/a_B} = 0,$$

откуда находим наиболее вероятное расстояние  $r_1 = a_B$ . ■

В этой простенькой задачке содержится важный факт: в основном состоянии электрон вовсе не вращается по орбите радиусом  $a_B$ . Если обращаться к классическим образам (а другие нам недоступны), электрон, скорее, можно представить себе как отрицательно заряженное облачко, плотность которого различна в разных точках пространства: в одних она выше, в других ниже. Вместо радиусов разрешенных орбит имеем теперь *области* наиболее вероятного местонахождения электрона (в химии их называют *орбиталями*). Соответственно из теории исчезли квантовые скачки с орбиты на орбиту, вместо этого мы имеем перестройку электронного облачка. А поскольку облачка — это визуальный образ плотности вероятности, их перестройка действительно может происходить мгновенно, так как никакие материальные объекты никуда не перемещаются.

Волновая функция  $\psi_{n00}(r)$  состояния с нулевым орбитальным моментом сферически симметрична, т. е. не зависит от углов. На рисунке 37.3 изображена плотность вероятности  $4\pi r^2 |\psi_{n00}|^2$  для трех нижних уровней с  $n = 1, 2, 3$ . Отчетливо видны расстояния, на которых найти электрон более вероятно, чем на других. Между ними имеются точки, где плотность вероятности равна нулю, — это узлы радиальной волновой функции, число которых на единицу меньше главного квантового числа  $n$ , т. е. равно максимальному значению азимутального квантового числа  $l$ .

**Задача 37.53.** Волновая функция основного состояния электрона в атоме водорода имеет вид уравнения (37.16). Найти вероятности  $W_1, W_2$  и  $W_3$  обнаружить электрон внутри сфер радиусами  $a_B, 2a_B$  и  $3a_B$ .



**Рис. 37.3.** Плотность вероятности найти электрон на расстоянии  $r$  от ядра для трех  $s$ -состояний ( $l = 0$ ) с главными квантовыми числами  $n = 1, 2, 3$ ; расстояния даны в единицах радиуса Бора

**Решение.** Вероятность найти электрон в элементе объема  $d\vec{r}$  равна  $dW = |\psi(r)|^2 d\vec{r}$ . Чтобы вычислить вероятность  $W(R)$  обнаружить электрон внутри сферы радиусом  $R$ , вероятность  $dW$  надо проинтегрировать по ограниченному этой сферой объему. Разница по сравнению с вычислением нормировки в задаче 37.51 лишь в том, что интеграл по радиальной переменной  $r$  надо брать от нуля до  $R$ , а не до бесконечности:

$$W(R) = \frac{4}{a_B^3} \int_0^R r^2 e^{-2r/a_B} dr = 1 - e^{-2R/a_B} \left[ 1 + 2\frac{R}{a_B} + 2\left(\frac{R}{a_B}\right)^2 \right]. \quad (37.19)$$

Отсюда находим:

$$\begin{aligned} W_1 &= W(R = a_B) = 1 - 5/e^2 = 0,323, \\ W_2 &= W(R = 2a_B) = 1 - 13/e^4 = 0,762, \\ W_3 &= W(R = 3a_B) = 1 - 25/e^6 = 0,938. \end{aligned} \quad (37.20)$$

Разность  $W_2 - W_1 \approx 0,439$  дает вероятность найти электрон *между* сферами с радиусами  $a_B$  и  $2a_B$ . Видно, что эта вероятность имеет тот же порядок величины, что и вероятность  $W_1$ . А вероятность найти электрон между сферами с радиусами  $2a_B$  и  $3a_B$  равна  $W_3 - W_2 = 0,176$ . ■

**Задача 37.54.** Найти напряженность электрического поля, создаваемого атомом водорода в основном состоянии. Во сколько раз электрическое поле, создаваемое атомом водорода в основном состоянии, слабее электрического поля его атомного ядра на расстояниях  $R = a_B$ ,  $R = 2a_B$  и  $R = 3a_B$ ?

**Решение.** Возьмем любую точку на расстоянии  $R$  от ядра. Мысленно проведем через нее сферу с центром на атомном ядре. По теореме Остроградского—Гаусса (см. т. 1,

разд. 15.8) поток вектора электрической напряженности через сферу создается зарядами, находящимися внутри нее. Это положительный заряд  $e$  протона и отрицательный заряд части электронного облачка, оказавшейся внутри воображаемой сферы. Поскольку в уравнении (37.19) вычислена вероятность  $W(R)$  нахождения электрона внутри сферы радиусом  $R$ , средний отрицательный заряд внутри этой сферы равен  $-eW(R)$ . Поэтому поле в точке  $R$ , создаваемое эффективным зарядом  $q(R) = e(1 - W(R))$ , имеет вид

$$E(R) = \frac{q(R)}{4\pi\epsilon_0 R^2} = \frac{e}{4\pi\epsilon_0 R^2} e^{-2R/a_B} \left[ 1 + 2\frac{R}{a_B} + 2\left(\frac{R}{a_B}\right)^2 \right]. \quad (37.21)$$

Поле атомного ядра водорода (протона) равно:

$$E_p(R) = \frac{e}{4\pi\epsilon_0 R^2}, \quad (37.22)$$

откуда имеем для отношения поля ядра к полю атома

$$\xi(R) = \frac{E_p(R)}{E(R)} = \frac{e^{2\rho}}{1 + 2\rho + 2\rho^2}, \quad \rho = \frac{R}{a_B}. \quad (37.23)$$

Отсюда  $\xi(R = a_B) = e^2/5 = 1,48$ ,  $\xi(R = 2a_B) = e^4/13 = 4,20$ ,  $\xi(R = 3a_B) = e^6/25 = 16,14$ . На больших расстояниях напряженность электрического поля (см. (37.21)) убывает экспоненциально, т.е. гораздо быстрее обычного кулоновского поля точечного заряда. Это так называемый эффект экранировки: отрицательный заряд электрона компенсирует положительный заряд ядра. При  $R \rightarrow 0$ , т.е. внутри электронного облачка, эффект экранировки ослабляется, и поле атома переходит в кулоновское поле точечного заряда ядра. ■

**Задача 37.55.** Электрон в атоме водорода находится в основном состоянии. Какова вероятность  $W_-$  обнаружить его в области, не доступной классической частице?

**Решение.** Область, доступная для классической частицы с энергией  $E < 0$ , удерживаемой кулоновским потенциалом, лежит в интервале радиальных расстояний  $0 \leq r \leq r_0$ . В точке поворота  $r_0$  полная энергия равна потенциальной энергии. Энергия в основном состоянии выражается уравнением (37.14), откуда получаем уравнение для точки поворота:

$$-\frac{e^2}{4\pi\epsilon_0 r_0} = -\frac{m_e e^4}{2(4\pi\epsilon_0)^2 \hbar^2}. \quad (37.24)$$

Отсюда следует, что

$$r_0 = 2\frac{4\pi\epsilon_0 \hbar^2}{m_e e^2} = 2a_B. \quad (37.25)$$

В задаче 37.53 мы уже вычислили вероятность  $W_2 = 0,762$  обнаружить электрон на расстояниях  $0 \leq r \leq r_0 = 2a_B$ . Тогда вероятность найти электрон в области  $r \geq r_0$  под кулоновским барьером равна  $W_- = 1 - W_2 = 0,238$ . В этой области плотность вероятности экспоненциально убывает. Это видно также из наших расчетов (см. задачу 37.53). Вероятность найти электрон под кулоновским барьером на расстоянии  $a_B$  от точки поворота  $r_0$  (т.е. между сферами радиусами  $2a_B$  и  $3a_B$ ) равна 0,176, а на все остальное подбарьерное пространство остается всего лишь вероятность  $1 - 0,938 = 0,062$ . Сумма вероятностей 0,176 и 0,062 как раз равна 0,238. Здесь мы снова сталкиваемся с возможностью квантовой частице (волне вероятности) проникнуть под потенциальный барьер, но на небольшое расстояние порядка ее длины волны де Бройля (т.е. порядка размеров атома). ■

Отличие квантовой механики от теории Бора — более богатая структура уровней: состояние определяется тремя квантовыми числами, как и в трехмерном потенциальном ящике. Кстати, это не случайно. Три квантовых числа в потенциальной яме и в атоме водорода — отражение трехмерности нашего пространства. Подсчитаем **кратность вырождения**, т.е. число различных состояний с одинаковой энергией (главным квантовым числом  $n$ ). При данном  $n$  число  $l$  пробегает значения от 0 до  $n - 1$  и каждому из них соответствует  $2l + 1$  значений  $m$ . Поэтому кратность вырождения  $N$  равна:

$$N = \sum_{l=0}^{n-1} (2l + 1) = n^2. \quad (37.26)$$

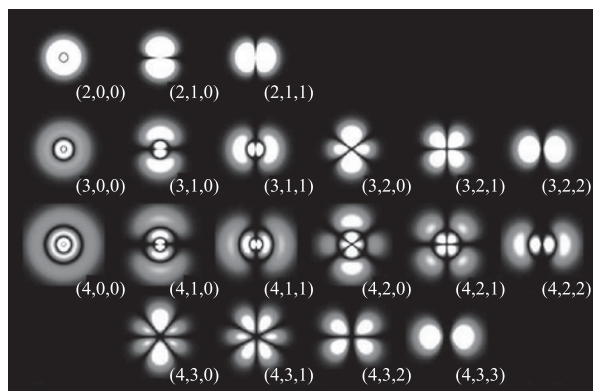
При  $n = 1$  имеем  $N = 1$ , т.е. основной уровень не вырожден. При  $n = 2$  кратность вырождения равна 4: уровень с  $l = 0$  и три уровня с  $l = 1$  и различными проекциями момента импульса  $m = -1, 0, +1$ . При  $n = 3$  кратность вырождения  $N = 9$ : один уровень с  $l = 0$ , три уровня с  $l = 1$  и пять уровней (по числу проекций) с  $l = 2$ . Для классификации уровней энергии по значению квантового числа  $l$  применяют условные обозначения, заимствованные из спектроскопии, где они появились еще до создания теории атома:

$l =$	0	1	2	3	4	5	...
символ	$s$	$p$	$d$	$f$	$g$	$h$	...

Главное квантовое число ставится впереди символа. Примеры возможных состояний (орбиталей):

$1s, 2s, 2p, 3s, 3p, 3d, 4s, 4p, 4d, 4f$  и т.д.

На рисунке 37.4 изображены первые орбитали атома водорода — поперечные сечения плотности вероятности, величина которой отражена яркостью (черный цвет соответствует минимальной плотности вероятности, а белый — максимальной). В левой колонке — сферически симметричные орбитали с  $l = 0$  и  $m = 0$ . Видны области максимальной вероятности найти электрон, а черные



**Рис. 37.4.** Несколько первых орбиталей атома водорода (в скобках указаны значения квантовых чисел  $n, l, m$ )

промежутки между ними указывают расстояния, где вероятность обнаружить его минимальна.

При переходе электрона с орбитали на орбиталь излучается фотон, уносящий собственный угловой момент (спин), равный  $\hbar$ . Следовательно, разрешены только переходы с изменением  $l$  на единицу: возникает **правило отбора**  $\Delta l = \pm 1$ . Это значит, что в атоме водорода допустимы переходы  $np \rightarrow 1s$ ,  $ns \rightarrow 2p$  и т. д., приводящие к тем же спектральным сериям, что и теория Бора. Более богатая структура уровней не проявляется пока в большом разнообразии атомных спектров из-за вырождения.

Говоря о вырождении уровней, мы имели в виду водородоподобный атом. В более сложных атомах или в присутствии внешних электромагнитных полей вырождение, как говорят, снимается и появляется зависимость энергии от чисел  $l, m$ . Любая некулоновская центрально-симметричная поправка к потенциальной энергии приведет к зависимости уровней энергии от  $l$  (наблюдается в щелочных металлах). В классической физике такая поправка к обычному закону притяжения (например, планеты к Солнцу) превращает эллиптические орбиты в незамкнутые кривые. Обращаясь по таким орбитам, планета движется по обычно эллипсу, который дополнительно вращается как целое, прецессирует в той же плоскости. Подобный эффект — вращение перигелия Меркурия — предсказала общая теория относительности. Новое движение — вращение орбиты как целого — приводит к дополнительной энергии, зависящей от углового момента.

Любое не центрально-симметричное поле (например, магнитное) снимет вырождение по магнитному квантовому числу  $m$  (вог где выясняется, что его незря назвали магнитным). В классической физике магнитное поле вызывает прецессию плоскости вращения вокруг направления поля и также появление из-за этого вращения дополнительной энергии. Сказанное можно сформулировать в виде общего вывода:

♦ **любое взаимодействие, нарушающее симметрию системы, снимает соответствующее ей вырождение уровней энергии.** ♦

В экспериментах снятие вырождения проявляется в расщеплении исходных спектральных линий на несколько компонентов.

## 37.4 Спин электрона и тонкая структура спектров

Дальнейшее исследование атомных спектров показало, что многие спектральные линии имеют два близких компонента. Так, еще в 1887 г. А. Майкельсон обнаружил расщепление  $\alpha$ -линии серии Бальмера в водороде (порождаемой переходом  $3p \rightarrow 2s$ ): она оказалась состоящей из двух линий со средней длиной волны  $6563 \text{ \AA}$  и разностью длин волн  $0,14 \text{ \AA}$  (т.е. относительная величина расщепления порядка  $10^{-5}$ ). Были обнаружены и линии, расщепленные на три, четыре и более компонентов. Расщепление линий, как мы теперь понимаем, означает расщепление энергетических уровней атома: с этим как раз и связана **тонкая структура спектральных линий**. Значит, существует неучтенное взаимодействие. Мы говорили, что расщепление линий появляется, например,



когда наложенное внешнее поле нарушает симметрию системы. А здесь неучтенное взаимодействие проявляется при отсутствии внешних полей, т. е. оно должно быть связано с какими-то внутренними свойствами атома.

Оказалось, что тут действительно проявились внутренние свойства, но не атома в целом, а электрона. Американские физики С. Гаудсмит и Дж. Уленбек выдвинули в 1925 г. гипотезу *спина электрона*: они предположили существование у электрона собственного момента импульса, не связанного с орбитальным движением. Сначала спин (от англ. *spin* — верчение) представляли себе как вращение электрона вокруг собственной оси, как аналог суточного вращения Земли. Потом осознали, что «верчение» нельзя понимать буквально: численные оценки давали линейную скорость верчения, превышающую скорость света. Сейчас модели спина не строят, принимая его за такое же неотъемлемое свойство, как масса или электрический заряд<sup>1</sup>.

◆ *Спин — собственный момент количества движения электрона, его врожденное квантовое свойство.* ◆

Существование спина остается загадкой, если находиться только в рамках квантовой механики Гейзенберга—Шрёдингера. Естественное объяснение спин получил только в релятивистской квантовой теории Дирака, соединившей теорию относительности с квантовой механикой.

Из опытов следовало, что электрону надо приписать *спиновое квантовое число*  $s = 1/2$ , имеющее те же свойства (см. формулу (37.9)), что и квантовое число  $l$ . Соответственно собственное значение оператора квадрата спина равно

$$\vec{S}^2 = \hbar^2 s(s+1) = \frac{3\hbar^2}{4}, \quad (37.27)$$

а проекция спина на какую-то ось (пробегаая через единицу  $\hbar$  все значения от максимального до минимального) выражается формулой

$$S_z = m_s \hbar, \quad (37.28)$$

где *магнитное спиновое число*  $m_s$  принимает в данном случае лишь два значения  $m_s = \pm 1/2$ .

Откуда же взялось расщепление спектральных линий? Попытаемся понять это с помощью полуклассических рассуждений. В классической физике любое вращение электрического заряда приводит к появлению магнитного поля. Вращающийся по орбите классический электрон можно представить как виток с током силой  $I$ , охватывающий площадь  $\Sigma$ , т. е. как диполь с магнитным моментом  $\mu = I\Sigma$  (эту формулу знал еще Ш. Кулон). Классическая оценка: электрон на

<sup>1</sup>Идеи вращения электрона вокруг собственной оси витали в воздухе и высказывались целым рядом авторов. Однако численные оценки вызывали недоумение и отторжение этой гипотезы. Гаудсмит и Уленбек тоже натолкнулись на эти же противоречия, но их научный руководитель П. Эрэнфест заявил, что они «достаточно молоды, чтобы позволить себе сделать одну глупость», и отослал их статью в печать. Возможно, неопределенность в установлении приоритета привела к тому, что введение в квантовую механику спина никак не было отмечено Нобелевским комитетом.

<sup>2</sup>Спиновое квантовое число для краткости принято называть спином, и мы в дальнейшем тоже будем использовать такую терминологию.

орбите радиусом  $R$  и скоростью  $v$  имеет период обращения  $T = 2\pi R/v$ . Возьмем какую-нибудь точку на орбите. За время  $T$  через нее проходит заряд  $e$ , т. е. сила тока по определению равна  $I = e/T = ev/(2\pi R)$ . Кроме того, электрон имеет орбитальный момент  $L = m_e v R$ , так что ток можно выразить через орбитальный момент, исключив скорость электрона:  $I = eL/(2\pi R^2 m_e) = eL/(2m_e \Sigma)$ , где мы ввели площадь «витка»  $\Sigma = \pi R^2$ . Тогда орбитальный магнитный момент, создаваемый электроном, равен  $\mu = eL/(2m_e)$ . Заменяем теперь в соответствии с правилами квантования  $L \rightarrow \hbar\sqrt{l(l+1)}$  и получим выражение для орбитального магнитного момента, которое может быть выведено и более строго:

$$\mu_L = \frac{e\hbar}{2m_e} \sqrt{l(l+1)}. \quad (37.29)$$

Из этих рассуждений следуют важные выводы.

- Естественная единица для магнитных моментов в микромире — так называемый **магнетон Бора**:

$$\mu_B = \frac{e\hbar}{2m_e} = 5,788 \cdot 10^{-5} \frac{\text{эВ}}{\text{Тл}} = 0,927 \cdot 10^{-23} \frac{\text{Дж}}{\text{Тл}}. \quad (37.30)$$

В этих единицах орбитальный магнитный момент (37.29) запишется как  $\mu_L = \mu_B \sqrt{l(l+1)}$ .

- Проекция магнитного момента на любую ось всегда должна быть целым кратным магнетона Бора:  $\mu_{Lz} = \mu_B m$ .
- Отношение орбитального магнитного момента  $\mu_L$  электрона к его механическому моменту импульса  $L$  равно  $e/(2m_e)$ ; его называют **гиромагнитным отношением**.

А вот к каким заключениям привели экспериментальные результаты.

- Спин электрона порождает **двойной магнетизм**: его собственный магнитный момент, связанный со спином, равен

$$\mu_s = 2\mu_B \sqrt{s(s+1)},$$

т. е. гиромагнитное отношение  $\mu_s/S = e/m_e$  оказалось для спина в 2 раза большим, чем для орбитального момента.

- Для проекции спинового магнитного момента свободного электрона на любую ось  $Oz$  имеем  $\mu_{sz} = 2\mu_B m_s = \pm\mu_B$ , поскольку  $m_s = \pm 1/2$ .

Удвоенное гиромагнитное отношение — лишнее доказательство, что электрон нельзя представлять как заряженный шарик, вращающийся вокруг собственной оси: в таком случае должно было бы получиться обычное гиромагнитное отношение. А для проекции спинового магнитного момента снова получились целые кратные магнетона Бора, как и в случае орбитального движения. По какой-то причине природа предпочитает иметь дело с целым магнетоном Бора, а не с его частями. Получается, что полуцелое значение собственного момента количества движения она компенсирует двойным гиромагнитным отношением.

Теперь можно понять, почему наличие у электрона собственного магнитного момента приводит к появлению какого-то неучтенного до сих пор взаимодействия. Для этого опять перейдем на полуклассический язык. Орбитальное движение электрона порождает магнитное поле, которое действует на собственный

магнитный момент электрона. Подобным образом магнитное поле Земли воздействует на стрелку компаса. Энергия этого взаимодействия сдвигает энергетические уровни атома, причем величина сдвига зависит, вообще говоря, от спинового и орбитального моментов количества движения.

**Важный вывод:** *взаимодействие спинового и орбитального магнитных моментов приводит к снятию вырождения и к расщеплению спектральных линий.*

**Задача 37.56.** Оценить расщепление уровней энергии вследствие взаимодействия магнитного спинового и магнитного орбитального моментов электрона в атоме водорода.

**Решение.** Круговой виток радиусом  $R$  с током силой  $I$  порождает в центре магнитное поле  $B = \mu_0 I / (2R)$ . В этой главе было показано, что вращающийся по орбите электрон можно представить как виток с током  $I = eL / (2\pi m R^2) \approx \mu_B / (\pi R^2)$ . Здесь для оценки мы приняли, что по порядку величины  $L \approx \hbar$ . Тогда получаем для магнитного поля, создаваемого орбитальным движением электрона в атоме, величину порядка  $B \approx \mu_0 \mu_B / (2\pi R^3)$ . Энергия взаимодействия собственного магнитного момента электрона с этим магнитным полем равна по порядку величины:

$$\Delta E \approx \mu_B B \approx \frac{\mu_0 \mu_B^2}{2\pi R^3}.$$

Для оценки полагаем  $R$  равным боровскому радиусу  $a_B$ . Подставляя сюда выражения для  $a_B$  и  $\mu_B$  и учитывая, что  $\mu_0 \epsilon_0 = 1/c^2$ , получаем оценку сдвига энергетических уровней:

$$\Delta E \approx \frac{mc^2}{2} \left( \frac{e^2}{4\pi\epsilon_0 \hbar c} \right)^4 = \frac{mc^2}{2} \alpha^4, \quad (37.31)$$

где  $\alpha$  — введенная выше (см. (35.6)) постоянная тонкой структуры. Энергия первого уровня атома водорода, как нам уже известно, равна  $E = -\alpha^2 mc^2 / 2$ , так что (37.31) можно переписать как  $\Delta E \approx \alpha^2 |E|$ . Поскольку  $\alpha \approx 1/137$ , а  $|E| = 13,6$  эВ, получаем  $\Delta E \approx 0,72$  мэВ, а относительный сдвиг уровней  $\delta E = \Delta E / |E| \approx \alpha^2 \approx 5 \cdot 10^{-5}$ , что соответствует экспериментальным данным. ■

Это и есть оценка (не расчет) искомого расщепления уровней. В сущности, расщепление уровней — это релятивистский эффект: по Бору скорость электрона на первой орбите  $v \approx \alpha c$  и  $\delta E \approx (v/c)^2$ . Поэтому до конца свойства спина могут быть поняты только в релятивистской квантовой теории. Мы не ставим себе такую задачу, а просто будем учитывать наличие у электрона этого удивительного свойства.

## Опыт Штерна—Герлаха

Прямое экспериментальное доказательство существования спина электрона было дано в опыте немецких физиков О. Штерна и В. Герлаха (1922). Идея опыта состоит в том, что в магнитном поле, не однородном по оси  $z$ , на электроны действует смещающая сила, направленная вдоль поля.

Происхождение этой силы проще уяснить сначала на примере электрического диполя, помещенного в электрическое поле. Электрический диполь представляет собой пару противоположных зарядов  $\pm e$ , расположенных на малом расстоянии  $l$  друг от друга. Величина электрического момента определяется как  $\vec{d} = e\vec{l}$ ,

причем вектор  $\vec{l}$  считается направленным от отрицательного заряда к положительному.

Пусть положительный заряд находится в точке  $\vec{r}$ , а отрицательный — в точке  $\vec{r}_1$ , так что  $\vec{r}_1 = \vec{r} - \vec{l}$ . Пусть диполь помещен в электрическое поле с напряженностью  $\vec{E}(\vec{r})$ . Найдем силу, действующую на диполь. На положительный заряд действует сила  $\vec{F}_1 = e\vec{E}(\vec{r})$ , на отрицательный —  $\vec{F}_2 = -e\vec{E}(\vec{r}_1)$ . Результирующая сила равна:

$$\vec{F} = \vec{F}_1 + \vec{F}_2 = e \left( \vec{E}(\vec{r}) - \vec{E}(\vec{r} - \vec{l}) \right).$$

Так как расстояние между зарядами мало, поле в точке расположения отрицательного заряда можно приближенно записать как

$$\vec{E}(\vec{r} - \vec{l}) \approx \vec{E}(\vec{r}) - (\vec{l} \cdot \vec{\nabla}) \vec{E}(\vec{r}).$$

Подставляя это разложение в выражение для силы  $\vec{F}$ , находим:

$$\vec{F} = e(\vec{l} \cdot \vec{\nabla}) \vec{E}(\vec{r}) = (\vec{d} \cdot \vec{\nabla}) \vec{E}(\vec{r}). \quad (37.32)$$

Если поле однородно ( $\vec{E}$  не зависит от  $\vec{r}$ ), то на заряды диполя действуют равные и противоположно направленные силы и результирующая сила равна нулю, как и следует из уравнения (37.32). Как известно, такая пара сил не смещает диполь (который в целом электрически нейтрален), а лишь поворачивает его вдоль поля (магнитный аналог — стрелка компаса). В неоднородном же поле результирующая сила отлична от нуля. В частном случае, когда поле  $\vec{E}$  зависит только от координаты  $z$ , в уравнении (37.32) отлична от нуля лишь производная по  $z$ :

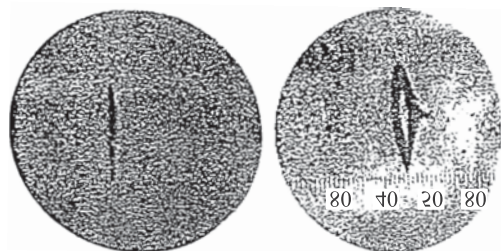
$$\vec{F} = d_z \frac{\partial \vec{E}(z)}{\partial z}, \quad (37.33)$$

где  $d_z$  — проекция электрического момента на ось  $z$ . Неоднородное поле стремится втянуть диполь в область, где оно сильнее.

Магнитных зарядов не существует, но магнитный диполь реализуется витком с током, и его свойства аналогичны свойствам электрического диполя. Поэтому в формуле (37.33) надо заменить электрическое поле на магнитное, электрический момент на магнитный и написать для силы, действующей на электрон в опыте Штерна—Герлаха, аналогичное выражение

$$\vec{F} = \mu_z \frac{\partial \vec{B}(z)}{\partial z}.$$

Схема опыта: пучок атомов пролетает сквозь неоднородное магнитное поле, направленное поперечно вектору скорости атомов. Сила, действующая на магнитные моменты атомов, отклоняет их. Соответственно возможным значениям проекции магнитного момента атомов на направление поля первоначальный пучок расщепляется на несколько пучков. Если полный магнитный момент атома определяется только спином электрона (как у атомов серебра в опыте Штерна—Герлаха), то первоначальный пучок расщепляется на два пучка (рис. 37.5). Для многоэлектронных атомов расщепленных пучков может быть больше. Для своего эксперимента Штерн и Герлах использовали серебро, которое испарялось в электрической печи. Численные значения расщепления составляли

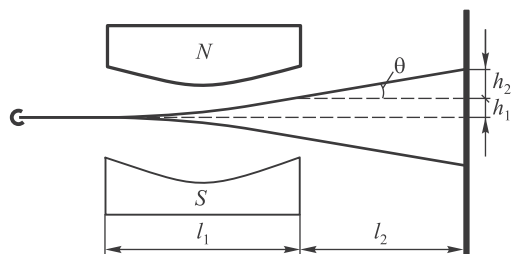


**Рис. 37.5.** Результат опытов Штерна—Герлаха (фото из их работы 1922 г.): пучок атомов серебра расщепился при прохождении неоднородного магнитного поля на два пучка, неотклоненного пучка не наблюдалось

доли миллиметра. Авторы подчеркнули в своих выводах, что неотклоненных атомов не было зарегистрировано.

Главный результат опытов Штерна—Герлаха — прямое экспериментальное доказательство *квантования проекции магнитного момента атомов*<sup>3</sup>. Согласно классической физике, первоначальный пучок должен не расщепиться, а размазаться в соответствии с произвольностью проекции магнитного момента на направление магнитного поля. Соответственно на экране за прибором вместо двух отдельных линий, оставленных атомами серебра, должна была бы наблюдаться размытая полоска.

**Задача 37.57.** Узкий пучок атомов со скоростью  $v$  и массой  $m$  пропускается через поперечное неоднородное магнитное поле, в котором на них действует сила  $F_z$ . Протяженность поля  $l_1$ , расстояние от магнита до экрана  $l_2$ . Определить отклонение следа пучка атомов на экране от его положения при отсутствии магнита (рис. 37.6).



**Рис. 37.6.** Отклонение атомов серебра неоднородным магнитным полем в опыте Штерна—Герлаха

**Решение.** Здесь мы имеем дело с задачей классической механики, которая позволяет подготовиться к количественному рассмотрению опыта Штерна—Герлаха. Время пролета атома через магнит равно  $t = l_1/v$ . Все это время на атом действует поперечная сила  $F_z$ , сообщающая ему поперечное ускорение  $a = F_z/m$ . За время пролета атом отклонится на расстояние  $h_1 = at^2/2 = F_z l_1^2 / (2mv^2)$  и приобретет поперечную скорость  $v_{\perp} = at = F_z l_1 / (mv)$ . Это значит, что из магнита атом вылетает под углом  $\theta$  к первоначальному направлению движения, причем  $\operatorname{tg} \theta = v_{\perp} / v$ . Следовательно, при пролете расстояния  $l_2$  до экрана атом отклонится еще и на расстояние  $h_2 = l_2 \operatorname{tg} \theta = l_2 v_{\perp} / v = F_z l_1 l_2 / (mv^2)$ . Складывая отклонения  $h_1$  и  $h_2$ , получаем искомое

<sup>3</sup>Заметим, что О. Штерн получил Нобелевскую премию (1943) «за вклад в развитие метода молекулярных пучков и за открытие магнитного момента протона», т.е. за совсем другие работы. Подтверждение существования спина и доказательство пространственного квантования направления магнитного момента атомов, равно как и сама гипотеза спина, признания Нобелевского комитета не нашла.

отклонение следа атома на экране:

$$h_{12} = \frac{F_z}{2mv^2} l_1(l_1 + 2l_2). \quad (37.34)$$

Часто в задаче стоит вопрос об отклонении пучка атомов при выходе из магнита. В таком случае надо полагать  $l_2 = 0$  в полученной формуле (37.34). ■

**Задача 37.58.** Узкий пучок атомов серебра при прохождении магнитного поля с неоднородностью  $\partial B/\partial z = 1$  кТл/м и протяженностью  $l_1 = 4$  см расщепился на два пучка. Экран для наблюдения установлен на расстоянии  $l_2 = 10$  см, скорость атомов  $v = 0,5$  км/с. Определить расстояние  $b$  между компонентами расщепленного пучка на экране.

**Решение.** Заполненные оболочки дают нулевой вклад в магнитный момент атома. Атом серебра имеет один валентный электрон, и потому его магнитный момент равен магнитному моменту электрона. В неоднородном поле на атом действует сила  $F_z = \pm \mu_B \frac{\partial B}{\partial z}$ , где  $\mu_B$  — магнетон Бора. Знаки « $\pm$ » соответствуют двум возможным направлениям магнитного момента электрона, и потому пучок расщепляется на два. На экране наблюдаются две полоски, находящиеся на расстоянии  $b$  друг от друга. Если в формулу (37.34) подставить выражение для силы  $F_z$ , то получим два отклонения  $s$ , различающиеся знаком. Поэтому искомое расщепление  $b = 2s$ . В результате приходим к выражению

$$b = \mu_B \left| \frac{\partial B}{\partial z} \right| \frac{l_1(l_1 + 2l_2)}{mv^2}. \quad (37.35)$$

Массу атома серебра найдем по таблице Менделеева: молярная масса серебра равна  $M = 107,868$  г/моль. Чтобы найти массу атома  $m$ , нужно  $M$  разделить на число Авогадро:  $m = 107,868 \cdot 10^{-3}/6,02 \cdot 10^{23} = 1,79 \cdot 10^{-25}$  кг. Подставим теперь в (37.35) численные значения:

$$b = 9,27 \cdot 10^{-24} \times 10^3 \times \frac{4 \cdot 10^{-2} \times (4 + 2 \times 10) \cdot 10^{-2}}{1,79 \cdot 10^{-25} \times (0,5 \cdot 10^3)^2} \approx 2 \cdot 10^{-3} \text{ м} = 2 \text{ мм}.$$

Такое расщепление вполне наблюдаемо в опытах. ■

Внимательный читатель непременно обратит внимание на то, что опыт Штерна—Герлаха, подтвердивший, как мы провозгласили, наличие спина у электрона, был проведен в 1922 г., а гипотеза спина получила признание лишь в 1926 г. Как же сами Штерн и Герлах интерпретировали свои результаты? То, что они категорически противоречат классической теории, было очевидно. Но авторы полагали, что они подтвердили гипотезу квантования орбитального момента. Теория атома Бора была усовершенствована Зоммерфельдом (1916), в ней появились новые квантовые числа, связанные с орбитальным моментом и его проекцией, которые квантовались по уже известным нам правилам. Однако орбитальный угловой момент не имеет ничего общего с результатами опыта Штерна—Герлаха: расщепление пучка серебра на два компонента четко указывает, что проекция магнитного момента атома принимает два значения, а орбитальный угловой момент всегда целочисленный и, значит, число расщепленных пучков должно быть больше двух. Кроме того, в этом случае обязательно должен был наблюдаться нерасщепленный пучок, соответствующий нулевому значению проекции углового орбитального момента. Введение в квантовую механику спина расставило все по своим местам.

### 37.5 Векторная модель атома

Обсудим качественные изменения, вносимые спином электрона в теорию атома. Полный момент импульса электрона  $\vec{J}$  складывается теперь из орбитальной  $\vec{L}$  и спиновой  $\vec{S}$  частей. Правила квантования полного момента такие же, как и для любых других моментов (орбитального и спинового): собственные значения квадрата оператора полного момента равны  $\vec{J}^2 = \hbar^2 j(j+1)$ . Возникает новое **квантовое число  $j$  полного момента импульса**, принимающее для электрона два значения  $j = l \pm \frac{1}{2}$  (при  $l = 0$  квантовое число  $j$  принимает одно значение  $j = 1/2$ ). Эти значения соответствуют двум случаям: спин параллелен и антипараллелен орбитальному моменту импульса. Необходимо ввести новые обозначения уровней, т. е. добавить индекс, указывающий величину полного момента импульса: уровни обозначают как  $nx_j$ , где  $n$  — главное квантовое число,  $x$  — прежний буквенный символ ( $s, p, d, f, \dots$ ) для обозначения величины азимутального квантового числа  $l$ . Как следствие нового вида взаимодействия, возникает более богатая структура атомных спектров, наблюдавшаяся на опыте. Проиллюстрируем это на примере первых возбужденных уровней атома водорода (табл. 37.1).

Энергия  $E$  уровней уже выражается не формулой Бора, а содержит поправки  $\Delta E$ , относительная величина которых равна  $\Delta E/E \approx \alpha^2$ . Мы не станем приводить эту формулу, но отметим ее характерное свойство: в отсутствие внешних полей энергия по-прежнему не зависит от квантового числа  $l$ , а определяется квантовым числом  $j$  полного момента импульса. Значит, к примеру, уровни  $2s_{1/2}$  и  $2p_{1/2}$  вырождены (их энергии совпадают), уровень же  $2p_{3/2}$ , как оказывается, лежит чуть выше.

Состояния многоэлектронных атомов классифицируются подобным образом. Если  $\vec{L}$  — суммарный орбитальный момент всех электронов, а  $\vec{S}$  — их суммарный спиновый момент, то полный момент системы определяется как  $\vec{J} = \vec{L} + \vec{S}$ . Соответственно и обозначается это состояние как  $^{2S+1}X_J$ . Под  $X$  понимается тот же буквенный символ, обозначающий значение орбитального момента количества движения (только в этом случае используется заглавная буква). Верхний левый

**Таблица 37.1.** Схема нижних уровней атома водорода

$n = 1,$	$l = 0,$	$j = \frac{1}{2} :$	$1s_{1/2}$
$n = 2,$	$l = \left\{ \begin{array}{l} 0, \\ 1, \end{array} \right.$	$j = \frac{1}{2} :$	$2s_{1/2}$
		$\left\{ \begin{array}{l} j = \frac{1}{2} : \\ j = \frac{3}{2} : \end{array} \right.$	$2p_{1/2}$ $2p_{3/2}$
$n = 3,$	$l = \left\{ \begin{array}{l} 0, \\ 1, \\ 2, \end{array} \right.$	$j = \frac{1}{2} :$	$3s_{1/2}$
		$\left\{ \begin{array}{l} j = \frac{1}{2} : \\ j = \frac{3}{2} : \end{array} \right.$	$3p_{1/2}$ $3p_{3/2}$
		$\left\{ \begin{array}{l} j = \frac{3}{2} : \\ j = \frac{5}{2} : \end{array} \right.$	$3d_{3/2}$ $3d_{5/2}$

индекс равен числу спиновых состояний (для одиночного электрона в нем не было необходимости, так как его спин всегда равен  $1/2$ ).

Итак, пусть дано состояние  $^{2S+1}X_J$ . Возникает вопрос: чему равен магнитный момент системы  $\vec{\mu}_J$ ? Ясно, что он направлен вдоль полного момента количества движения  $\vec{J}$ , а его размерность определяется магнетоном Бора  $\mu_B$ . Тогда

$$\begin{aligned}\vec{\mu}_J &= g\mu_B \frac{\vec{J}}{\hbar}, \\ \mu_J &= g\mu_B \sqrt{J(J+1)}.\end{aligned}\quad (37.36)$$

Для гиромагнитного отношения (обобщение аналогичной величины, связанной с орбитальным и спиновым моментами) можно тогда написать выражение вида  $g \cdot e/(2m_e)$ . Коэффициент пропорциональности  $g$  называется **множителем Ланде** или просто  **$g$ -фактором**. Для орбитального движения  $g = 1$ , для спинового магнитного момента  $g = 2$ . Задача о магнитном моменте атома сводится к нахождению зависимости  $g$  от квантовых чисел  $J$ ,  $L$  и  $S$ .

Ответ можно получить с помощью простой полуклассической модели, получившей название **векторной модели атома**. Сначала возведем в квадрат уравнение, связывающее  $\vec{J}$  с  $\vec{L}$  и  $\vec{S}$ :  $\vec{J}^2 = \vec{L}^2 + 2\vec{L}\vec{S} + \vec{S}^2$ . Скалярное произведение векторов орбитального момента и спина можно записать тогда как  $\vec{L}\vec{S} = (\vec{J}^2 - \vec{L}^2 - \vec{S}^2)/2$ . Далее квадраты моментов можно выразить через соответствующие квантовые числа по уже известным нам правилам. Находим тогда выражение для скалярного произведения:

$$\vec{L}\vec{S} = \frac{\hbar^2}{2} (J(J+1) - L(L+1) - S(S+1)). \quad (37.37)$$

Полный магнитный момент складывается из магнитного момента, создаваемого суммарным орбитальным моментом количества движения, и суммарного спинового магнитного момента. Но спин, как уже говорилось, обладает двойным магнетизмом. Поэтому с учетом уравнения (37.36) можно записать:

$$g\mu_B \frac{\vec{J}}{\hbar} = \mu_B \frac{\vec{L}}{\hbar} + 2\mu_B \frac{\vec{S}}{\hbar}.$$

Сокращая общий множитель  $\mu_B$  и умножая обе части на  $\vec{J}/\hbar$  (в правой части  $\vec{J}$  заменен на  $\vec{L} + \vec{S}$ ), получаем

$$gJ(J+1) = \frac{1}{\hbar^2} (\vec{L} + 2\vec{S})(\vec{L} + \vec{S}) = L(L+1) + 2S(S+1) + \frac{3}{\hbar^2} (\vec{L}\vec{S}).$$

Если подставить сюда выражение (37.37) для скалярного произведения  $\vec{L}\vec{S}$ , то получим окончательный ответ:

$$g = 1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)}. \quad (37.38)$$

Убедимся, что эта формула воспроизводит уже известные результаты. Если полный спиновый момент равен нулю, то полный момент совпадает с орбитальным. Подставляя в (37.38) значения  $S = 0$ ,  $J = L$ , получаем  $g = 1$ , как и должно быть для магнитного момента, создаваемого чисто орбитальным движением



электронов. В обратном случае нулю равен орбитальный момент и полный момент количества движения равен спиновому. Подставляя в (37.38) значения  $L = 0, J = S$ , находим  $g = 2$  в полном согласии с двойным магнетизмом спинового момента. Именно такой случай реализуется для элементов первой группы в опыте Штерна—Герлаха.

Упоминалось, что для сложных атомов (например, серы) расщепление пучков будет более сложным. Теперь мы можем предсказать результат количественно. Основное состояние серы  ${}^3P_2$ , т. е.  $S = 1, L = 1, J = 2$ . Из формулы (37.38) для множителя Ланде легко получаем  $g = 3/2$ , так что магнитный момент атома равен:

$$\vec{\mu}_J = \frac{3}{2} \mu_B \frac{\vec{J}}{\hbar}.$$

Проекция магнитного момента на ось  $z$  (так же как и проекция  $m_J$  полного углового момента) принимает пять различных значений в соответствии с правилами квантования момента:

$$\mu_{Jz} = \frac{3}{2} \mu_B m_J, \quad m_J = 0, \pm 1, \pm 2.$$

Теперь ясно, что пучок атомов серы в опыте Штерна—Герлаха расщепится на пять компонентов, причем одна из них ( $m_J = 0$ ) не будет отклоняться магнитным полем.

## 37.6 Принцип Паули и валентность элементов

Мы видели, что орбитальный момент импульса характеризуется азимутальным квантовым числом  $l$ , принимающим целые значения. Собственный спиновый момент характеризуется спиновым квантовым числом  $s$ , равным  $1/2$  для электрона. Спин — фундаментальное квантовое свойство всех элементарных частиц. В соответствии со значением спинового квантового числа  $s$  частицы делятся на два класса: **бозоны** (целые  $s$ ) и **фермионы** (полуцелые  $s$ ). Другие значения  $s$  науке неизвестны. Примеры фермионов: электрон, мюон, нейтрино, кварки, частицы из трех кварков (протон, нейтрон). Примеры бозонов: фотон, частицы из двух кварков ( $\pi$ -мезоны и т. д.), промежуточные векторные бозоны  $W^\pm$  и  $Z$ , ответственные за перенос слабых ядерных взаимодействий, открытый в 2012 г. бозон Хиггса (об элементарных частицах мы поговорим более подробно в главе 45).

Бозоны и фермионы — это не имена собственные каких-то частиц, а названия целых семейств. Бозоны и фермионы подчиняются соответственно *квантовой статистике Бозе—Эйнштейна* и *Ферми—Дирака*, которые мы рассмотрим позднее, в части VII учебника. Каждая микрочастица принадлежит одному из этих семейств и поменять принадлежность уже не в состоянии. В теории элементарных частиц замечено, что из фермионов построено вещество, а бозоны переносят фундаментальные взаимодействия.

Если собрать вместе несколько бозонов, то они не будут мешать друг другу занять низшее энергетическое состояние и, следовательно, все там окажутся.

Поведение коллектива фермионов регулируется **принципом Паули**:

◆ **два одинаковых фермиона не могут находиться в одном и том же квантовом состоянии.** ◆

В применении к атому принцип Паули запрещает двум электронам иметь одинаковый набор квантовых чисел: состояния электронов должны отличаться значением хотя бы одного из них. Если бы принципа Паули не существовало, то у всех атомов в основном состоянии все электроны находились бы на нижнем энергетическом уровне в соответствии с принципом минимума энергии системы, и атомы различных элементов были бы до ужаса похожи друг на друга. Таким образом, известная нам химия, все разнообразие элементов и их свойств (в частности, гамма вкусовых ощущений, отличающая, скажем, вино от сыра) — следствия принципа Паули. В 1945 г. швейцарский физик В. Паули был удостоен Нобелевской премии «за открытие принципа запрета, названного его именем».

До сих пор мы в основном рассматривали только атом с одним электроном. Станем добавлять лишний электрон и соответственно увеличивать на единицу заряд ядра. Иными словами, совершим прогулку по таблице Менделеева. Введем понятия *оболочки* как совокупности уровней с одинаковыми главными квантовыми числами  $n$  и *подоболочки* (уровни данной оболочки с одинаковыми  $l$ ). Традиционные обозначения для оболочек:

$$\begin{array}{l} n = \quad 1, \quad 2, \quad 3, \quad 4, \quad \dots \\ \text{символ} \quad K, \quad L, \quad M, \quad N, \quad \dots \end{array}$$

Число уровней в оболочке мы фактически уже подсчитали: оно равно кратности вырождения по  $m$  и  $l$ , умноженной на 2 в соответствии с двумя возможными направлениями спина электрона. Таким образом, оболочка может вместить  $2n^2$  электронов, а подоболочка —  $2(2l + 1)$  электронов (число различных значений  $m$ , удвоенное по указанной выше причине). Состояния в атоме различаются четырьмя квантовыми числами  $n, l, m, m_s$ . По принципу Паули лишь один электрон может иметь данный конкретный набор квантовых чисел.

Напомним: три квантовых числа  $n$ ,  $l$  и  $m$  отражают трехмерность пространства,  $m_s$  — внутренние свойства электрона.

Порядок заполнения состояний в многоэлектронных атомах определяется принципом минимума энергии уровней с учетом влияния уже заполненных оболочек. Нижние (более близкие к ядру) оболочки частично экранируют заряд ядра, что ведет к некулоновскому полю для высших оболочек. Этим объясняется зависимость энергии уровней в многоэлектронных атомах от квантового числа  $l$ . Чем больше  $l$ , тем дальше находится электрон от ядра и тем слабее его связь с ядром, т. е. такой уровень лежит выше. При больших моментах (состояния  $d$ ,  $f$  и т. п.) уровень близок к  $s$ -подуровням следующей оболочки или даже выше их. В целом заполнение идет в таком порядке (случаются небольшие отклонения):

$$1s \parallel 2s, 2p \parallel 3s, 3p \parallel 4s, 3d, 4p \parallel 5s, 4d, 5p \parallel 6s, 4f, 5d, 6p \parallel \text{и т. д.}$$

В этом ряду знаком  $\parallel$  отделены периоды таблицы Менделеева. В соответствии с полученными формулами для числа уровней в зависимости от  $n$  и  $l$  1-й период содержит 2 элемента, 2-й и 3-й — по 8, а 4-й и 5-й — по 18 элементов. Здесь «вклиниваются» заполнения  $d$ -подоболочек (по 10 элементов — переходные металлы). В 6-м периоде добавляется заполнение  $f$ -подоболочки (14 элементов — лантаниды) и т. д.

Химические свойства зависят в основном от строения внешней электронной оболочки (валентных электронов). Отсюда и вытекает периодичность повторения

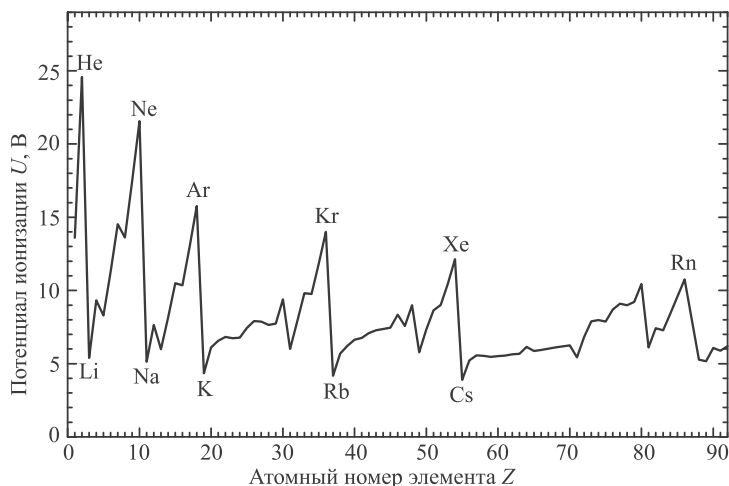


Рис. 37.7. Потенциал ионизации атомов химических элементов от водорода до урана

свойств элементов. У благородных газов, замыкающих периоды таблицы Менделеева, заполнены внешние  $s$ - и  $p$ -подоболочки из восьми электронов. У элементов первой группы (щелочные металлы) на внешней оболочке имеется всего один электрон, а галогенам (VII группа) не хватает одного электрона для заполнения внешней  $p$ -подоболочки. Все эти рассуждения хорошо знакомы по курсу химии, но они носят качественный характер. Как известно, валентность элемента определяется не столько количеством электронов, сколько степенью легкости, с которой их можно вырвать из атома. Понятно, что внешние электроны оторвать легче, но хотелось бы получить хоть какой-нибудь количественный критерий. Таким критерием может служить **потенциал ионизации** — энергия, которую необходимо затратить, чтобы удалить электрон из атома. Расчет многоэлектронных атомов непрост, но в целом таблица Менделеева объясняется квантовой механикой. Мы ограничимся лишь самыми простыми атомами и точные расчеты заменим количественными оценками.

Экспериментальные результаты по величине потенциала ионизации атомов химических элементов от водорода до урана представлены на рис. 37.7. Ясно видна периодичность свойств элементов: пики потенциала ионизации наблюдаются для благородных газов, минимумы — для щелочных металлов.

## Эффективный заряд ядра, оценки потенциала ионизации и закон Мозли

Мы в деталях изучили энергетические уровни водородоподобных атомов, описываемые формулой Бора

$$E_n = -13,6 \text{ эВ} \frac{Z^2}{n^2}.$$

Для более сложных систем со многими электронами эта формула неверна, но мы будем ее использовать, вводя поправку на экранировку ядра электронами

путем замены заряда ядра  $Z$  на некий *эффективный* заряд  $Z_{\text{эф}}$ . Большой точности при сравнении с экспериментальными данными так не добиться, но мы удовлетворимся совпадением по порядку величины.

**Водород.** Заряд ядра  $Z = 1$ , экранировки нет, в основном состоянии  $n = 1$ . Энергия основного уровня  $E_1 = -13,6$  эВ, минимальная энергия электрона, вырванного из атома, равна нулю. Следовательно, ионизационный потенциал  $U$  — это абсолютное значение энергии основного состояния:  $U = 13,6$  эВ. Это значение будет служить нам ориентиром, позволяющим не заблудиться в энергетической шкале микромира.

**Положительный ион гелия.** Ион  $\text{He}^+$  отличается от водорода только удвоенным зарядом ядра:  $Z = 2$ . Отсюда  $U = 13,6 \times 4 = 54,4$  эВ.

**Гелий.** Нейтральный атом гелия в основном состоянии имеет два электрона на нижней оболочке ( $n = 1$ ), различающиеся проекциями спинов. Представим себе такую картину. Когда второй электрон находится дальше от ядра, чем первый, то заряд ядра от него экранируется и равен (с его «точки зрения») единице. Когда же второй электрон располагается к ядру ближе первого, то «видимый» им заряд ядра равен двум. Оба электрона равноправны, так что описанные ситуации равновероятны. Поэтому для оценки эффективного заряда ядра мы возьмем среднее арифметическое:  $Z_{\text{эф}} = (1+2)/2 = 3/2$ . Конечно, электрон не может полностью экранировать ядро от своего партнера. Конечно, надо принять во внимание энергию кулоновского отталкивания электронов. И все же получающаяся оценка не столь плоха:

$$U \approx 13,6 \frac{Z_{\text{эф}}^2}{n^2} = 13,6 \frac{9}{4} = 30,6 \text{ эВ.}$$

Эксперимент дает  $U = 24,6$  эВ. Бросается в глаза огромная величина ионизационного потенциала у гелия (самый большой в таблице Менделеева). Инертность гелия — прямое следствие этого факта.

**Дважды ионизованный атом лития  $\text{Li}^{++}$ .** Водородоподобная система с  $Z = 3$ . Поэтому  $U = 13,6 \times 9 = 122,4$  эВ.

**Ион лития.** Ион  $\text{Li}^+$  подобен атому гелия, но для него эффективный заряд ядра на единицу больше:  $Z_{\text{эф}} = 5/2$ . Отсюда  $U \approx 13,6 \times (25/4) = 85$  эВ. Эксперимент дает  $U = 81$  эВ.

**Литий.** Третий электрон в нейтральном атоме лития располагается на второй оболочке — уровне с  $n = 2$ . По этой причине два внутренних электрона почти полностью экранируют от него две единицы заряда ядра:  $Z_{\text{эф}} = 1$ . Отсюда  $U \approx 13,6/4 = 3,4$  эВ. Эксперимент дает  $U = 5,4$  эВ.

Эти оценки очень показательны: насколько легче (в сравнении с водородом) оторвать от лития один электрон и как трудно удалить последующие. Поэтому литий одновалентен.

**Бериллий.** Ионы  $Be^{+++}$  и  $Be^{++}$  аналогичны водороду и гелию, но из-за больших зарядов ядер ( $Z_{эф} = 4$  и  $Z_{эф} = 7/2$  соответственно) имеют огромные потенциалы ионизации. Ион  $Be^+$  аналогичен атому водорода: внутренние электроны экранируют две единицы заряда ядра, так что для внешнего электрона  $Z_{эф} = 2$ , но он находится на второй оболочке с  $n = 2$ , и поэтому потенциал ионизации получается для него таким же, как для атома водорода:  $U = 13,6$  эВ. Эксперимент дает 18,2 эВ. В нейтральном атоме бериллия на второй оболочке помещаются два электрона. Система похожа на атом гелия: для внешних электронов эффективный заряд ядра равен  $Z_{эф} \approx 3/2$ , но они находятся на второй оболочке ( $n = 2$ ), что в 4 раза уменьшает потенциал ионизации по сравнению с нашим значением для гелия. В итоге получаем оценку  $U = 30,6/4 = 7,8$  эВ, экспериментальное же значение  $U = 10,4$  эВ.

Отсюда вывод: поскольку первые два электрона вырвать из атома бериллия гораздо легче, чем последующие, то Be двухвалентен.

Понятие эффективного заряда ядра полезно также при рассмотрении свойств так называемого *характеристического рентгеновского излучения*, возникающего при переходе внешних электронов на свободное место на внутренних оболочках. Как мы выяснили, для электронов на  $K$ -оболочке  $Z_{эф} \approx Z - 1/2$ . Например, для меди  $Z = 29$  и  $Z_{эф} \approx 28,5$ . При переходе внешних, далеких от ядра электронов с энергией, почти равной нулю, испускается фотон с энергией  $E = 13,6 \times (28,5)^2 = 11\,000$  эВ. Длина волны такого фотона  $\lambda = c/\nu = hc/E = 1,12 \text{ \AA}$ .

В 1913 г. был установлен **закон Мозли**, связывающий частоту  $\nu$  характеристического рентгеновского излучения элемента и его атомный номер  $Z$ :

$$\boxed{\sqrt{\frac{\nu}{c R_y}} = \frac{Z - \sigma_n}{n}}, \quad (37.39)$$

где  $R_y$  — постоянная Ридберга,  $c$  — скорость света,  $n$  — главное квантовое число оболочки, на которую совершается переход, а  $\sigma_n$  — некая постоянная. В этом законе легко теперь увидеть проявление экранирования ядра, т. е. влияние на отдельный электрон атома всех остальных электронов. Исторически закон Мозли окончательно подтвердил, что свойства элемента зависят от атомного номера  $Z$ , а не от атомной массы. Это устранило последние сомнения в правильности понимания расположения элементов в периодической системе.

## Электронная конфигурация атомов

Для атома любого элемента мы можем указать его электронную конфигурацию. Теперь познакомимся с двумя правилами Хунда, позволяющими узнать для основного состояния каждого атома значение его суммарных моментов: спинового  $S$ , орбитального  $L$  и полного  $J$  (здесь мы, как договорились ранее, будем использовать термин «момент», имея в виду соответствующее квантовое число).

- *Наименьшей энергией обладает атом с наибольшим возможным при данной электронной конфигурации значением  $S$  и наибольшим возможным при этом  $S$  значением  $L$ .*

- Если в незаполненной подоболочке атома находится не более половины максимально возможного для нее числа электронов, то  $J = |L - S|$ . Если же подоболочка заполнена более чем наполовину, то  $J = L + S$ .

Правила Хунда являются эмпирическими (т.е. не выведенными из теории, а установленными экспериментально); второе правило выводится из первого и из полученной выше формулы (37.37) для скалярного произведения  $\vec{L}\vec{S}$ . Для облегчения применения правил Хунда полезно запомнить, что:

- 1) надо рассматривать только незаполненную электронную подоболочку, так как моменты электронов в заполненных подоболочках взаимно компенсируются;
- 2) значения  $S$  и  $L$  одинаковы для двух подоболочек, из которых одна имеет столько электронов, сколько не хватает для заполнения другой.

Для примера рассмотрим атом серы. Ее атомный номер  $Z = 16$ , и, следовательно, электронная конфигурация имеет вид  $1s^2 2s^2 2p^6 3s^2 3p^4$ . Индексы справа сверху обозначают числа электронов в указанных состояниях. Видим, что незаполненной остается третья оболочка, на которой находятся четыре электрона с орбитальными моментами, равными единице. При данном  $L = 1$  согласно принципу Паули параллельными могут быть лишь спины трех электронов, различающихся значениями  $m = -1, 0, 1$ . Спин четвертого электрона должен быть направлен в противоположную сторону, и потому максимально возможное значение суммарного спина для данной конфигурации равно  $S = 3/2 - 1/2 = 1$ .

При этом значении  $S$  сумма проекций орбитальных моментов электронов с одинаково направленными спинами равна нулю, следовательно, максимально возможное значение проекции суммарного орбитального момента определяется четвертым электроном и равно единице. Поэтому  $L = 1$ . Так как внешняя подоболочка заполнена более чем наполовину, то  $J = L + S = 2$ . Мы пришли к выводу, что основным состоянием атома серы является  ${}^3P_2$ .

Для сравнения найдем основное состояние атома кремния ( $Z = 14$ ). От атома серы он отличается тем, что на внешней подоболочке у него находятся два, а не четыре  $p$ -электрона. Для ее заполнения не хватает четырех электронов, т.е. столько электронов, сколько их имеется у серы. Поэтому для атома кремния получаем те же значения  $S = 1$  и  $L = 1$ . Внешняя подоболочка заполнена менее чем наполовину, откуда  $J = |L - S| = 0$ . Стало быть, основным состоянием кремния является  ${}^3P_0$ .

## 37.7 Принцип тождественности частиц

Две частицы тождественны, если все их физические свойства в точности совпадают, что исключает возможность экспериментально различать их. В классической теории всегда предполагается, что мы можем в принципе проследить за движением даже одинаковых частиц и сказать, какая из них куда полетела. Поэтому в классической теории одинаковые частицы принципиально различимы. В квантовой механике это не так: принцип неопределенности не позволяет проследивать траектории и, стало быть, неразличимость частиц имеет принципиальный характер и влияет на результаты вычислений.

Пусть, например, система из двух частиц описывается гамильтонианом (оператором энергии)  $\hat{H}(\vec{r}_1, \vec{r}_2)$  и пусть система находится в состоянии с волновой

функцией  $\Psi(\vec{r}_1, \vec{r}_2)$ . Введем оператор  $\hat{P}_{12}$ , переставляющий частицы местами, т. е. обменивающий их радиусы-векторы:

$$\hat{P}_{12}\Psi(\vec{r}_1, \vec{r}_2) = \Psi(\vec{r}_2, \vec{r}_1). \quad (37.40)$$

Математически тождественность частиц можно записать как условие коммутации операторов  $\hat{H}$  и  $\hat{P}_{12}$ :

$$\hat{P}_{12}\hat{H} = \hat{H}\hat{P}_{12}. \quad (37.41)$$

Это условие обеспечивает физическую неразличимость частиц, поскольку тогда волновая функция  $\Psi(\vec{r}_2, \vec{r}_1)$  будет также решением уравнения Шрёдингера с тем же значением энергии. Действительно, если  $\hat{H}\Psi(\vec{r}_1, \vec{r}_2) = E\Psi(\vec{r}_1, \vec{r}_2)$  и мы подействуем на обе части этого уравнения оператором  $\hat{P}_{12}$ , то получим  $\hat{P}_{12}\hat{H}\Psi(\vec{r}_1, \vec{r}_2) = E\Psi(\vec{r}_2, \vec{r}_1)$ . Из-за условия коммутации мы можем пронести оператор  $\hat{P}_{12}$  через гамильтониан:  $\hat{P}_{12}\hat{H}\Psi(\vec{r}_1, \vec{r}_2) = \hat{H}\hat{P}_{12}\Psi(\vec{r}_1, \vec{r}_2) = \hat{H}\Psi(\vec{r}_2, \vec{r}_1)$ , и наше уравнение примет вид  $\hat{H}\Psi(\vec{r}_2, \vec{r}_1) = E\Psi(\vec{r}_2, \vec{r}_1)$ .

Но мы помним, что коммутирование операторов с гамильтонианом означает сохранение их собственного значения. Найдем собственные значения  $p$  оператора  $\hat{P}_{12}$ . Для этого надо решить уравнение

$$\hat{P}_{12}\Psi(\vec{r}_1, \vec{r}_2) = p\Psi(\vec{r}_1, \vec{r}_2). \quad (37.42)$$

С учетом определения (37.40) оператора  $\hat{P}_{12}$  записываем это уравнение в виде

$$\Psi(\vec{r}_2, \vec{r}_1) = p\Psi(\vec{r}_1, \vec{r}_2). \quad (37.43)$$

Снова подействуем на обе части (37.43) оператором перестановки частиц:

$$\Psi(\vec{r}_1, \vec{r}_2) = p\Psi(\vec{r}_2, \vec{r}_1). \quad (37.44)$$

Из (37.43) и (37.44) получаем, что

$$\Psi(\vec{r}_1, \vec{r}_2) = p^2\Psi(\vec{r}_1, \vec{r}_2), \quad (37.45)$$

т. е.  $p = \pm 1$ . Таким образом, оператор  $\hat{P}_{12}$  может иметь только два собственных значения. При  $p = 1$  волновая функция симметрична относительно операции перестановки частиц:  $\Psi(\vec{r}_2, \vec{r}_1) = \Psi(\vec{r}_1, \vec{r}_2)$ . При  $p = -1$  имеем антисимметричную волновую функцию:  $\Psi(\vec{r}_2, \vec{r}_1) = -\Psi(\vec{r}_1, \vec{r}_2)$ . Таким образом, мы получили важный результат:

◆ *волновые функции системы двух тождественных частиц могут быть либо симметричными, либо антисимметричными относительно операции перестановки частиц местами.* ◆

Обобщение этого результата: состояния системы тождественных частиц либо симметричны, либо антисимметричны относительно перестановки любых двух из них. Какое состояние реализуется — зависит от природы рассматриваемых частиц. Частицы с симметричными состояниями называются **бозонами**, с антисимметричными — **фермионами**. Ранее так мы называли частицы с целым и полуцелым спинами соответственно, но никакой ошибки мы не делали. В

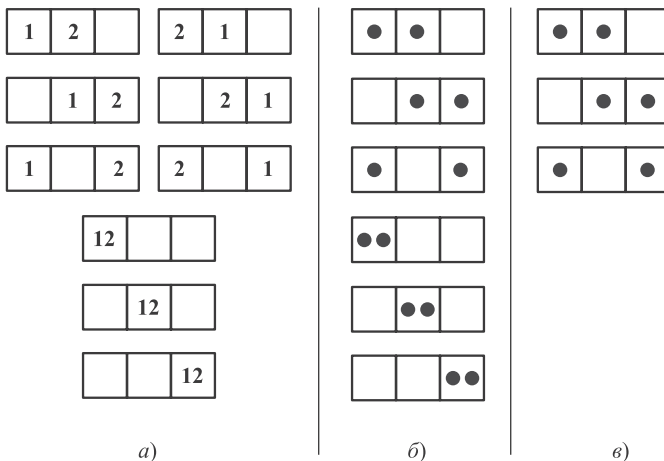
релятивистском уравнении Дирака, в отличие от уравнения Шрёдингера, спин частиц возникает автоматически. Существует фундаментальная **теорема о связи спина со статистикой** (В. Паули, 1940): частицы с полуцелым спином ( $s = 1/2, 3/2$  и т. д.) описываются антисимметричными волновыми функциями, а с целым ( $s = 0, 1, 2$  и т. д.) — симметричными. Теорема Паули является следствием объединения квантовой механики с теорией относительности.

Обратимся для примера к состоянию двух атомных электронов. В пренебрежении их взаимодействием волновая функция распадается на произведение одночастичных волновых функций каждого электрона:  $\Psi(\vec{r}_1, \vec{r}_2) = \Psi_i(\vec{r}_1)\Psi_j(\vec{r}_2)$ . Индексы  $i, j$  обозначают здесь полный набор квантовых чисел ( $n, l, m, s$ ), которыми одно состояние отличается от другого. Меняя электроны местами, приходим к состоянию с той же энергией, описываемому волновой функцией  $\Psi_i(\vec{r}_2)\Psi_j(\vec{r}_1)$ . Поэтому в силу принципа суперпозиции возможны состояния, описываемые любой линейной комбинацией этих двух функций, причем все они будут иметь ту же энергию. Из теоремы Паули следует, что для электронов со спином  $s = 1/2$  физический смысл имеет лишь антисимметричная комбинация

$$\Psi(\vec{r}_1, \vec{r}_2) = C[\Psi_i(\vec{r}_1)\Psi_j(\vec{r}_2) - \Psi_i(\vec{r}_2)\Psi_j(\vec{r}_1)]. \quad (37.46)$$

Если состояния электронов одинаковы ( $i = j$ , т. е. совпадают **все** квантовые числа), то  $\Psi(\vec{r}_1, \vec{r}_2) = 0$ . Мы снова пришли к принципу Паули: не может быть двух электронов в одном состоянии.

Принцип тождественности (неразличимости) частиц оказывает колоссальное влияние на законы распределения микрочастиц по возможным состояниям, т. е. на **квантовую статистику**. На рисунке 37.8 показано число способов распределения двух частиц по трем состояниям (показаны клеточками). Классические частицы можно условно перенумеровать, и ситуации, когда они меняются местами, будут принципиально различимы. Всего, таким образом, имеется девять



**Рис. 37.8.** Число способов распределения двух частиц по трем состояниям для классических частиц (а), бозонов (б) и фермионов (в)



способов распределения двух частиц по трем клеточкам (рис. 37.8, а). Для квантовых частиц состояния с переставленными местами частицами неразличимы, и остается всего шесть способов размещения двух бозонов по трем клеточкам (рис. 37.8, б). Из них три способа соответствуют двум частицам в одном состоянии (клеточке). Для фермионов этот вариант исключен принципом запрета Паули, и у двух фермионов остается всего три способа размещения по трем состояниям (рис. 37.8, в). Ясно, что при большом числе частиц свойства ансамбля бозонов и фермионов будут резко различаться. Но об этом мы поговорим в следующих главах курса.

## Контрольные вопросы

1. Сформулируйте правила квантования момента импульса и проекции момента на выделенную ось.
2. Какими квантовыми числами описывается состояние электрона в атоме?
3. От каких квантовых чисел зависит энергия электрона в атоме водорода?
4. Какова кратность вырождения энергетического уровня в атоме водорода?
5. Какова размерность волновой функции электрона в СИ?
6. Чему равен потенциал ионизации атома водорода? Какова зависимость энергии электрона в атоме водорода от главного квантового числа?
7. В каких аспектах квантовая теория атома совпадает с полуклассической теорией Бора, а в каких эти теории различны?
8. Перечислите свойства электрона, связанные с наличием у него спина.
9. Что такое магнетон Бора? Чему равно гиромагнитное отношение для орбитального движения и для спина?
10. Опишите схему опыта Штерна—Герлаха. От чего зависит число компонентов, на которые расщепляется пучок атомов в этом опыте?
11. Почему направление магнитного момента электрона противоположно направлению его спина?
12. Сформулируйте принцип Паули. Что такое электронные оболочки и подоболочки в атоме? Объясните роль принципа Паули в повторяемости (периодичности) свойств химических элементов.
13. Как связан потенциал ионизации с валентностью элемента? Почему электрон труднее удалить из атома неона, чем из атома натрия?
14. Предположим, что никакого спина у электрона не существует, так что его состояние в атоме определяется лишь тремя квантовыми числами  $n, l, m$ . Какие последствия это имело бы для таблицы элементов Менделеева и нашей жизни?
15. Предположим, что спин электрона равен  $3/2$ . Какой элемент играл бы тогда роль гелия — первого из группы благородных газов? Какой элемент был бы вторым в этой же группе?
16. Найдите электронную конфигурацию основного состояния атома фосфора ( $Z = 15$ ).
17. На каких физических предпосылках основан принцип неразличимости тождественных частиц и каковы его последствия?
18. Каковы свойства симметрии волновых функций системы тождественных частиц относительно операции перестановки любых двух из них?
19. Сформулируйте теорему о связи спина со статистикой.

## Глава 38

# Физическая природа химической связи

### 38.1 Молекулы

Атомы представляют собой связанные электрическими силами системы ядер и электронов. В свою очередь, атомы под действием тех же сил в определенных условиях могут объединяться в молекулы. Первоначально химия предполагала существование специфических «химических сил» для объяснения образования молекул из атомов, а также химических реакций между атомами или молекулами. Однако таких «химических сил» не существует: образование молекул обусловлено обычными электрическими (кулоновскими) силами взаимодействия заряженных частиц, т. е. электронов и ядер, из которых состоят атомы. Но механизм образования молекул можно понять только на основе квантовой механики. Для простоты ограничимся рассмотрением простейших двухатомных молекул. Силы, удерживающие атомы в молекуле, вызваны взаимодействием внешних электронов, а электроны внутренних оболочек при объединении атомов в молекулы остаются в прежних состояниях. Различают два вида связи между атомами в молекуле: *ионную (гетерополярную)* и *ковалентную (гомеополярную)*.

#### Ионная связь

Этот вид химической связи обусловлен переносом валентных электронов с одного атома на другой и кулоновским притяжением образовавшихся ионов. Иными словами, *ионная связь* осуществляется в том случае, когда около одних атомных ядер в молекуле образуется избыток электронов, а около других — их недостаток. Тогда молекулу можно представить как образование, состоящее из отрицательного и положительного ионов, притягивающихся друг к другу (например, молекула NaCl состоит из ионов  $\text{Na}^+$  и  $\text{Cl}^-$ ).

Для описания ионных молекул еще до появления квантовой механики с успехом применялись полуэмпирические способы расчета энергии связи, основанные на классической электростатике. Пусть, например, мы имеем дело с молекулами типа KCl, NaCl и т. п. Обозначим через  $U_m$  ионизационный потенциал щелочного

металла, а через  $U_h$  средство галогена к электрону. В начальном состоянии имеем два нейтральных атома (например, К и Cl), разнесенные на бесконечно большое расстояние. Попытаемся мысленно построить из них молекулу. Затрачивая энергию  $U_m$ , мы отрываем электрон от металла и переносим его на галоген. При присоединении электрона к последнему выделяется энергия  $U_h$ , так что энерговыделение на этом этапе составляет  $U_h - U_m$ . Далее сближаем образованные нами ионы до равновесного расстояния  $R_0$  (размер молекулы). При этом выделяется дополнительная энергия кулоновского взаимодействия  $E_C$ . Полная выделившаяся энергия и есть энергия связи  $E_{св}$  получившейся молекулы; именно ее надо затратить, чтобы разорвать молекулу на составные части:

$$E_{св} = U_h - U_m + E_C. \quad (38.1)$$

Учитывая, что кулоновская энергия  $E_C = e^2/(4\pi\epsilon_0 R_0)$ , находим отсюда выражение для равновесного расстояния:

$$R_0 = \frac{e^2}{4\pi\epsilon_0(E_{св} - U_h + U_m)}. \quad (38.2)$$

Применим полученную формулу для оценки размера молекулы KCl. Берем экспериментальные данные:  $U_m = 4,3$  эВ,  $U_h = 3,8$  эВ,  $E_{св} = 4,8$  эВ, так что кулоновская энергия

$$E_C = 4,8 - 3,8 + 4,3 = 5,3 \text{ эВ} = 5,3 \times 1,6 \cdot 10^{-19} = 8,5 \cdot 10^{-19} \text{ Дж.}$$

Отсюда легко находим размер молекулы:

$$R_0 = \frac{(1,6 \cdot 10^{-19})^2}{4\pi \times 8,854 \cdot 10^{-12} \times 8,5 \cdot 10^{-19}} = 2,7 \cdot 10^{-10} \text{ м} = 2,7 \text{ \AA}. \quad (38.3)$$

Получившаяся величина примерно в 5 раз превышает радиус Бора и дает вполне приемлемую оценку размера молекулы.

Оценим также размер молекулы другого соединения — HCl, для которого  $U_m = 13,6$  эВ и  $E_{св} = 2,7$  эВ. Отсюда находим энергию кулоновского взаимодействия:  $E_C = 2,7 - 3,8 + 13,6 = 12,5$  эВ. Это в  $12,5/5,3 = 2,36$  раза больше кулоновской энергии для молекулы KCl. Соответственно размер молекулы будет в такое же число раз меньшим:  $R_0 = 2,7 \text{ \AA}/2,36 = 1,1 \text{ \AA}$  (эксперимент дает близкое значение  $R_0 = 1,3 \text{ \AA}$ ).

Отчуждение электрона одним ионом у другого (ионная связь) встречается в большинстве неорганических соединений, особенно в молекулах из атомов далеких групп. Кроме соединений щелочных металлов и галогенов, можно привести и другие примеры. Однако изложенные классические соображения не дают возможности вычислить независимо энергию связи молекулы (или ее размер). Само возникновение равновесного расстояния остается вне пределов компетенции такой теории. Силы электростатического притяжения между ионами должны уравниваться в равновесном состоянии какими-то силами отталкивания. Природа этих сил довольно сложна и связана с перекрытием оболочек ионов. При сближении атомов оболочки начинают сильно деформироваться, что препятствует дальнейшему сближению. Величина  $R_0$  — это равновесное расстояние между ионами, на котором силы электрического притяжения уравниваются квантово-механическим отталкиванием атомов на близких расстояниях.

## Ковалентная связь

Другой вид связи — **ковалентная связь** — наблюдается в молекулах, где электроны, осуществляющие связь, значительную часть времени проводят в пространстве между атомами и в какой-то мере являются «общими» для обоих ядер. Гомеополлярная (ковалентная) связь характерна для большинства молекул с двумя одинаковыми атомами ( $\text{H}_2$ ,  $\text{N}_2$ ,  $\text{O}_2$  и т.п.). Гомеополлярная связь не поддается классическому описанию, а требует квантового рассмотрения.

Рассмотрим простейшую гомеополлярную молекулу — молекулу водорода. Впервые такую задачу решили в 1927 г. немецкие физики В. Гайтлер и Ф. Лондон, и эта работа заложила основу нового научного направления — квантовой химии. Мы ограничимся принципиальным анализом, опуская проведение расчетов, так как нашей задачей является уяснение физической природы ковалентной (гомеополлярной) связи. Схема взаимодействий в такой молекуле, состоящей из двух протонов (ядер атома водорода)  $A$  и  $B$  и двух электронов 1 и 2, показана на рис. 38.1.

Уравнение Шрёдингера для системы имеет вид

$$\hat{H}\Psi(\vec{r}_1, \vec{r}_2) = E\Psi(\vec{r}_1, \vec{r}_2), \quad (38.4)$$

где  $\vec{r}_1$  и  $\vec{r}_2$  — координаты первого и второго электронов, а  $\hat{H}$  — гамильтониан системы:

$$\begin{aligned} \hat{H} &= \hat{H}_1 + \hat{H}_2 + \hat{U}, \\ \hat{H}_1 &= -\frac{\hbar^2}{2m_e} \nabla_1^2 - \frac{e^2}{4\pi\epsilon_0 r_{1A}}, \quad \hat{H}_2 = -\frac{\hbar^2}{2m_e} \nabla_2^2 - \frac{e^2}{4\pi\epsilon_0 r_{2B}}, \\ \hat{U} &= -\frac{e^2}{4\pi\epsilon_0 r_{1B}} - \frac{e^2}{4\pi\epsilon_0 r_{2A}} + \frac{e^2}{4\pi\epsilon_0 r_{12}} + \frac{e^2}{4\pi\epsilon_0 R}. \end{aligned} \quad (38.5)$$

Гамильтониан  $\hat{H}_1$  содержит оператор кинетической энергии электрона 1 и потенциальную энергию того же электрона в кулоновском поле ядра  $A$ . Аналогичную структуру имеет гамильтониан  $\hat{H}_2$ . Оператор  $\hat{U}$  описывает четыре взаимодействия: электрона 1 с ядром  $B$ , электрона 2 с ядром  $A$ , электронов друг с другом и ядер также друг с другом. Ядра имеют массу, примерно в 2000 раз превышающую массу электрона, поэтому они движутся гораздо медленнее электронов, и в первом приближении их можно считать неподвижными (такое

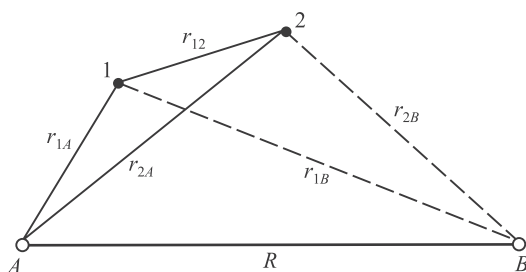


Рис. 38.1. Схема взаимодействий в молекуле водорода

приближение называется адиабатическим). Поэтому волновая функция  $\Psi(\vec{r}_1, \vec{r}_2)$  рассматривается как функция координат только электронов, а расстояние  $R$  между ядрами, важное в ковалентной связи, входит в волновую функцию  $\Psi$  как параметр. Получающиеся из уравнения Шрёдингера собственные значения энергии будут тогда зависеть от расстояния  $R$ , т. е.  $E = E(R)$ , причем в случаях параллельной и антипараллельной ориентации спинов электронов характер этой зависимости оказывается различным.

Полная волновая функция зависит не только от пространственных координат электронов, но и от их спинов. В соответствии с принципом Паули (согласно которому данный набор квантовых чисел может иметь лишь одна частица) такая полная волновая функция должна быть антисимметричной относительно перестановки электронов. Поскольку спиновое состояние не зависит от орбитального, пространственные и спиновые переменные разделяются, и полную волновую функцию электронов в молекуле водорода можно представить в виде произведения координатной и спиновой волновых функций:

$$\Psi_{12} = \Psi(\vec{r}_1, \vec{r}_2)\Phi(\vec{s}_1, \vec{s}_2). \quad (38.6)$$

Из принципа Паули следует, что:

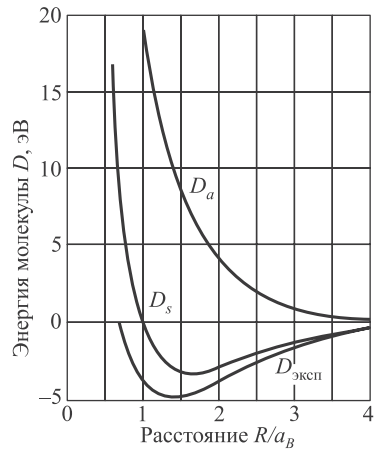
- 1) если спиновая волновая функция симметрична, то пространственная функция антисимметрична;
- 2) если спиновая волновая функция антисимметрична, то пространственная функция должна быть симметричной.

Симметричная спиновая волновая функция соответствует параллельному расположению спинов электронов, т. е. суммарному спину электронов  $S = 1$ . Такое состояние называется *триплетным* по числу  $2S + 1 = 3$  возможных проекций суммарного спина. Во втором же случае антисимметричная спиновая функция связана с противоположно ориентированными спинами, дающими в сумме  $S = 0$ , и, следовательно, возможно только одно значение проекции полного спина. Такое состояние называется *синглетным*.

Взаимодействие нейтральных атомов, когда расстояние между ядрами  $R$  много больше характерных размеров электронных оболочек (т. е. радиуса Бора  $a_B = 4\pi\epsilon_0\hbar^2/(me^2) \approx 5,3 \cdot 10^{-11}$  м), практически отсутствует, а энергия системы  $E$  равна удвоенной энергии основного состояния атома водорода  $2E_0$ , где  $E_0 = -Ry = -13,6$  эВ. При уменьшении  $R$  до нуля энергия системы будет бесконечно возрастать из-за кулоновского отталкивания ядер. Связанное состояние молекулы будет существовать, если между этими двумя асимптотическими значениями  $R$  имеются расстояния, для которых  $E < 2E_0$ . Равновесным будет тогда состояние с наименьшим из этих значений  $E = E_{\min}$ . Разность  $D = E_{\min} - 2E_0$  называется *энергией молекулы*. Если она отрицательна, то для развала молекулы надо затратить энергию  $|D|$  — *энергию диссоциации молекулы*. Стало быть, такое состояние устойчиво. Если же  $D > 0$ , то такая молекула существовать не может, она распадется на атомы сама по себе, без наших дополнительных усилий.

Результаты расчетов можно описать следующим образом. Если спины электронов параллельны, то при сближении атомов водорода энергия системы монотонно увеличивается, никакого связанного состояния не возникает (кривая

**Рис. 38.2.** Энергия  $D$  молекулы водорода (в эВ) в зависимости от расстояния (в единицах радиуса Бора) между ядрами. Верхняя кривая — теоретический расчет для энергии  $D_a$  триплетного ( $S = 1$ ) состояния с антисимметричной координатной волновой функцией. Связанное состояние молекулы в этом случае отсутствует. Средняя кривая — теоретический расчет для энергии  $D_s$  синглетного ( $S = 0$ ) состояния с симметричной координатной волновой функцией. Нижняя кривая — экспериментальный результат: равновесное расстояние между ядрами порядка полутора радиусов Бора, энергия молекулы  $D_{\text{эксп}} = -4,48$  эВ (рисунок из книги: П. Гомбаши. Проблема многих частиц в квантовой механике. М.: ИЛ, 1953. С. 106)



$D_a$  на рис. 38.2). В этом случае предоставленные самим себе два атома вновь разойдутся, стремясь к состоянию с наименьшей возможной энергией. Условие существования связанного состояния — наличие минимума энергии системы при некотором значении  $R_0$  — выполняется только при синглетном состоянии электронов, когда их спины антипараллельны (кривая  $D_s$  на рис. 38.2).

Понять такое поведение энергии можно на качественном уровне, если рассмотреть два водородных атома, находящихся на большом расстоянии  $R \gg a_B$  друг от друга. Тогда членом взаимодействия  $\hat{U}$  в гамильтониане (см. (38.5)) можно пренебречь, и мы имеем два независимых атома, описываемых суммой гамильтонианов  $\hat{H}_1$  и  $\hat{H}_2$ . Волновые функции их основных состояний мы обозначим здесь через  $\psi_0$ . Из них можно построить решение для суммы гамильтонианов  $\hat{H}_1 + \hat{H}_2$ . Оно будет представлено в виде произведения волновых функций  $\psi_0(\vec{r}_{1A})\psi_0(\vec{r}_{2B})$ , а энергия — в виде суммы энергий, т.е. будет равна  $2E_0$ . Физический смысл произведения волновых функций очевиден: электрон 1 находится в поле ядра  $A$ , а электрон 2 — в поле ядра  $B$ . Но электроны неразличимы, и эта физическая ситуация ничем не отличается от случая, когда электрон 2 находится в поле ядра  $A$ , а электрон 1 — в поле ядра  $B$ . Этому соответствует иное произведение волновых функций:  $\psi_0(\vec{r}_{2A})\psi_0(\vec{r}_{1B})$ , которое тоже является решением уравнения Шрёдингера с той же энергией  $2E_0$ . Согласно принципу суперпозиции, решением с этой же энергией будет и любая линейная комбинация указанных произведений. Поскольку принцип Паули требует, чтобы полная волновая функция была либо симметрична, либо антисимметрична по пространственным координатам электрона, она должна иметь вид

$$\Psi_{\pm}(\vec{r}_1, \vec{r}_2) = N_{\pm} [\psi_0(\vec{r}_{1A})\psi_0(\vec{r}_{2B}) \pm \psi_0(\vec{r}_{2A})\psi_0(\vec{r}_{1B})], \quad (38.7)$$

где  $N_{\pm}$  — вычисляемые нормировочные коэффициенты, обеспечивающие равенство единице полной вероятности:  $\int \Psi_{\pm}^2 d\vec{r}_1 d\vec{r}_2 = 1$ . Подставляя сюда выражения (38.7) и учитывая, что  $\psi_A(\vec{r})$  и  $\psi_B(\vec{r})$  уже нормированы на единицу, получаем:

$$N_{\pm}^2 (1 \pm 2P^2 + 1) = 1, \quad (38.8)$$

откуда

$$N_{\pm} = \frac{1}{\sqrt{2(1 \pm P^2)}}. \quad (38.9)$$

Величина  $P$ , называемая *интегралом перекрытия*, равна

$$P = \int \psi_0(\vec{r}_{1A})\psi_0(\vec{r}_{1B}) d^3\vec{r}_1. \quad (38.10)$$

Она характеризует степень перекрытия волновых функций электронов, сосредоточенных вблизи ядер  $A$  и  $B$ , и играет важную роль в теоретических расчетах свойств молекул. Иными словами, это численное выражение степени независимости друг от друга пространственного распределения двух электронов.

Если учесть теперь член взаимодействия  $\hat{U}$ , то в зависимости от симметрии волновой функции  $\Psi_{\pm}$  получатся поправки к энергии разных знаков. В случае антисимметричного состояния  $\Psi_{-}$  (полный спин электронов  $S = 1$ ) поправка получается положительной (отталкивание атомов) и связанного состояния не возникает. В случае же симметричного состояния  $\Psi_{+}$  (полный спин электронов  $S = 0$ ) поправка для больших  $R$  отрицательна (притяжение атомов), что приводит к образованию молекулы водорода.

Теоретически вычисленная энергия молекулы имеет следующую структуру:

$$D_s = \frac{Q + A}{1 + P^2}, \quad D_a = \frac{Q - A}{1 - P^2}. \quad (38.11)$$

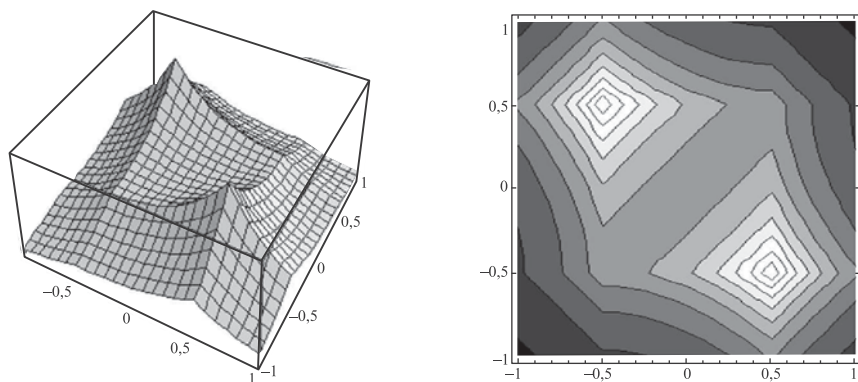
Выражение  $Q$  представляет собой сумму слагаемых вида (множители  $e^2/(4\pi\epsilon_0)$  мы здесь и далее опускаем для упрощения записи):

$$\begin{aligned} & \int \frac{\psi_0^2(\vec{r}_{1A})\psi_0^2(\vec{r}_{2B})}{r_{12}} d^3\vec{r}_{1A} d^3\vec{r}_{2B}; \\ & \int \frac{\psi_0^2(\vec{r}_{1A})}{r_{1B}} d^3\vec{r}_{1A}, \quad \int \frac{\psi_0^2(\vec{r}_{2B})}{r_{2A}} d^3\vec{r}_{2B}; \\ & \frac{1}{R}. \end{aligned} \quad (38.12)$$

Слагаемое в первой строке характеризует кулоновское взаимодействие двух электронов (расстояние  $r_{12}$  в знаменателе), каждый из которых «размазан» в пространстве с плотностью  $\psi_0^2$ . Первое выражение во второй строке описывает кулоновское взаимодействие с ядром  $B$  ( $r_{1B}$  в знаменателе) отрицательно заряженного электронного облачка, находящегося у ядра  $A$  (аргумент  $\vec{r}_{1A}$  у волновой функции). Аналогично, второе выражение в той же строке соответствует взаимодействию с ядром  $A$  электрона, «размазанного» вокруг ядра  $B$ . Наконец, выражение в третьей строке, не зависящее от волновых функций электронов, описывает кулоновское взаимодействие атомных ядер, находящихся на расстоянии  $R$  друг от друга.

Гораздо интереснее *обменная энергия*  $A$ . Она также является суммой нескольких слагаемых:

$$\begin{aligned} & \int \frac{\psi_0(\vec{r}_{1A})\psi_0(\vec{r}_{1B})}{r_{1A}} d^3\vec{r}_{1A}; \\ & \int \frac{\psi_0(\vec{r}_{1A})\psi_0(\vec{r}_{1B})}{r_{1B}} d^3\vec{r}_{1A}; \\ & \int \frac{\psi_0(\vec{r}_{1A})\psi_0(\vec{r}_{2A})\psi_0(\vec{r}_{1B})\psi_0(\vec{r}_{2B})}{r_{12}} d^3\vec{r}_{1A} d^3\vec{r}_{2B}. \end{aligned} \quad (38.13)$$



**Рис. 38.3.** Плотность электронного облака в молекуле водорода в синглетном ( $S = 0$ ) состоянии, соответствующем симметричной координатной части волновой функции (слева) и контурные графики (справа)

В этих слагаемых нет квадратов волновых функций, описывающих связь электрона с ядром, зато есть произведения волновых функций разных аргументов. Слагаемое в первой строке выглядит как притяжение к ядру  $A$  электрона, который «принадлежит» обоим ядрам. Аналогичный смысл имеет слагаемое во второй строке. Наконец, слагаемое в третьей строке описывает отталкивание двух электронов, которые будто бы обмениваются своими состояниями.

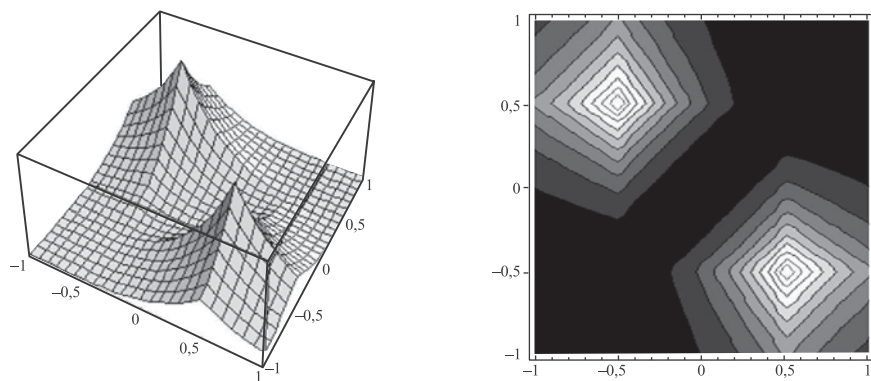
Осмыслить поведение энергии диссоциации помогают трехмерные изображения плотности электронов (квадратов волновых функций) и соответствующие им контурные графики (аналогичные изогипсам на топографических картах).

На рисунке 38.3 изображена электронная плотность в синглетном ( $S = 0$ ) состоянии, в котором координатная волновая функция симметрична относительно перестановки частиц. Расстояния показаны в условных единицах, в которых координаты ядра  $A$  равны  $\{-0,5; 0,5; 0\}$ , а ядра  $B$  —  $\{0,5; -0,5; 0\}$ . Видно, что с наибольшей вероятностью электроны могут быть обнаружены вблизи ядер. Но симметричная координатная волновая функция заметно отлична от нуля и в области между ядрами. В центре между атомами возникает электронное облачко из двух электронов. Они притягивают к себе протоны — возникает стягивающая сила, преодолевающая из-за близости облачка к ядрам кулоновское отталкивание между ними. В результате и образуется молекула водорода.

На рисунке 38.4 показана электронная плотность в триплетном ( $S = 1$ ) состоянии с антисимметричной координатной волновой функцией. И здесь с наибольшей вероятностью электроны могут быть обнаружены вблизи ядер, но между ними распределение плотности электронов совершенно иное: электроны стремятся расположиться подальше друг от друга — в центре системы волновая функция равна нулю. Отталкивание протонов ничем не экранируется и, более того, усугубляется притяжением электронов, расположившихся в периферических областях системы. Это приводит к увеличению энергии системы и невозможности антисимметричного связанного состояния.

Таким образом, химическая связь молекулы водорода возникает из-за двух эффектов. Во-первых, каждый электрон не принадлежит какому-то одному





**Рис. 38.4.** Плотность электронного облака в триплетном ( $S = 1$ ) состоянии, соответствующем антисимметричной координатной части волновой функции (слева) и контурные графики (справа)

ядру, а «размазан» вокруг обоих ядер: он как бы «прыгает» от одного протона к другому, образуя стягивающее облачко в пространстве между ними. Если расстояния между протонами велики, этот процесс затруднен. Если расстояния малы, то у электрона остается меньше пространства между протонами и в соответствии с соотношением неопределенностей растет его кинетическая энергия. Следовательно, существует некоторое равновесное расстояние между протонами, при котором энергия системы минимальна. Этот механизм работает и в ионе водорода  $\text{H}_2^+$ , состоящем из двух протонов и электрона. Но стягивающий эффект у одного электрона меньше, чем у двух, откуда и разница в энергии молекулы:  $D = -2,65$  эВ у иона водорода против  $D = -4,48$  эВ у молекулы водорода.

Помимо описанного механизма в молекуле водорода действует и другой эффект, отсутствующий в ионе водорода. Здесь возникает *обменное взаимодействие*, связанное с симметрией волновой функции относительно перестановки электронов местами. Во избежание недоразумений следует подчеркнуть, что квантовая механика не вводит никаких новых взаимодействий.

♦ *Обменное взаимодействие — это часть известных нам электростатических взаимодействий, отличающихся тем, что их вклад в энергию системы зависит от свойств симметрии координатных волновых функций относительно перестановок частиц. Обменное взаимодействие — это чисто квантовый эффект, не имеющий аналогов в классической теории электромагнетизма.* ♦

## Комбинации различных типов связи

Мы познакомились с двумя типами связи — ионной и ковалентной. В первой основную роль играет «перескок» электронов от одного атома к другому, во второй — дополнительные силы притяжения между атомами, порождаемые парой электронов с противоположными спинами. На самом деле оказывается, что такое резкое разделение не есть внутреннее свойство системы, а, скорее, недостаток наших методов расчета. Так, для улучшения согласия с экспериментом использованное приближение

Гайтлера—Лондона может быть уточнено. Именно, ранее исходным моментом были волновые функции электронов — электрона 1 в поле ядра  $A$  и электрона 2 в поле ядра  $B$ . Но электрон 1 также находится и в поле ядра  $B$ , а электрон 2 — в поле ядра  $A$ . Эти взаимодействия были учтены в поправочных членах, входящих в оператор  $\hat{U}$ . Можно, однако, попытаться включить эти взаимодействия в рассмотрение с самого начала. Для этого представим волновую функцию первого электрона в виде

$$\psi_1(\vec{r}_1) = c\psi_A(\vec{r}_1) + d\psi_B(\vec{r}_1),$$

где  $c, d$  — некие численные коэффициенты. Они не являются независимыми, условие нормировки волновой функции  $\psi_1$  приводит к связи между ними:

$$\int \psi_1^2(\vec{r}_1) d^3\vec{r}_1 = c^2 + d^2 + 2cdP = 1, \quad (38.14)$$

где  $P$  — тот же интеграл перекрытия (см. (38.10)). Аналогичный вид имеет и волновая функция второго электрона:

$$\psi_2(\vec{r}_2) = c\psi_A(\vec{r}_2) + d\psi_B(\vec{r}_2).$$

Волновую функцию синглетного двухэлектронного состояния строим как произведение:

$$\begin{aligned} \Psi_+(\vec{r}_1, \vec{r}_2) = \psi_1(\vec{r}_1)\psi_2(\vec{r}_2) = & c^2\psi_A(\vec{r}_1)\psi_A(\vec{r}_2) + d^2\psi_B(\vec{r}_1)\psi_B(\vec{r}_2) + \\ & + cd[\psi_A(\vec{r}_1)\psi_B(\vec{r}_2) + \psi_B(\vec{r}_1)\psi_A(\vec{r}_2)]. \end{aligned} \quad (38.15)$$

Мы видим, что здесь содержится волновая функция Гайтлера—Лондона — это слагаемое во второй строке выражения (38.15). Однако в приведенных формулах содержится и описание другого случая: это слагаемые  $\psi_A(\vec{r}_1)\psi_A(\vec{r}_2)$  и  $\psi_B(\vec{r}_1)\psi_B(\vec{r}_2)$ . Физический смысл первого из них достаточно очевиден: оба электрона сосредоточены вокруг ядра  $A$ , образуя ион водорода  $H^-$  в то время, как ядро  $B$ , лишившееся электрона, образует ион  $H^+$ . Мы видим, что этот случай описывает ионную связь. Второе слагаемое описывает аналогичную ситуацию, только ядра  $A$  и  $B$  меняются ролями. Для симметричной молекулы при обмене ядер местами ничего не меняется, откуда следует равенство коэффициентов  $c^2 = d^2 = |cd|$ .

Таким образом, волновая функция (38.15) описывает некую комбинацию ионной и ковалентной связей, причем относительные веса этих двух типов связи одинаковы. Физически это, однако, не так, поскольку мы уже убедились, что метод Гайтлера—Лондона дает неплохое описание молекулы водорода. Но мы нащупали путь. Ясно, что недостаток метода Гайтлера—Лондона заключается в пренебрежении ионными членами, а недостаток только что изложенного подхода (его называют *методом молекулярных орбиталей*) — в переоценке влияния этих членов.

Можно принять оба этих метода как два различных исходных приближения, и естественным обобщением может стать улучшение численных значений коэффициентов в волновой функции, определяющих статистический вес ионной и ковалентной связей. Представим, например, симметричную волновую функцию как суперпозицию отдельных двухэлектронных состояний с произвольными коэффициентами  $a_1, a_2, a_3$ :

$$\begin{aligned} \Psi_+(\vec{r}_1, \vec{r}_2) = & a_1\psi_A(\vec{r}_1)\psi_A(\vec{r}_2) + a_2\psi_B(\vec{r}_1)\psi_B(\vec{r}_2) + \\ & + a_3[\psi_A(\vec{r}_1)\psi_B(\vec{r}_2) + \psi_B(\vec{r}_1)\psi_A(\vec{r}_2)]. \end{aligned} \quad (38.16)$$

При использовании такой волновой функции согласие с экспериментом заметно улучшается. В реальных расчетах коэффициенты  $a_1, a_2, a_3$  определяются из условия минимума средней энергии, и их численные значения получаются автоматически, в зависимости

от свойств системы. Если окажется, что численное значение одного из них намного превышает по абсолютной величине два других, то можно говорить о преобладании связи определенного типа. Так, если  $|a_1| \gg |a_2|, |a_3|$ , то это будет ионная связь  $A^-B^+$  (электрон «перепрыгнул» с ядра  $B$  на ядро  $A$ ); если  $|a_2| \gg |a_1|, |a_3|$ , то это случай ионной связи  $A^+B^-$  (электрон перешел с ядра  $A$  на ядро  $B$ ); если же, наконец,  $|a_3| \gg |a_1|, |a_2|$ , то мы имеем дело с ковалентной связью.

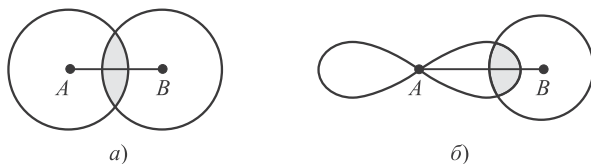
Этот подход применим как к молекулам из двух разных атомов, так и к соединениям одинаковых атомов и позволяет учесть вклад ковалентного соединения в ионных молекулах и вклад ионной связи в ковалентных молекулах. В применении к симметричным молекулам должно выполняться равенство  $|a_1| = |a_2|$ , но третий коэффициент  $a_3$  остается независимым, в отличие от метода молекулярных орбиталей. Сравнение экспериментов с расчетами по волновой функции (см. (38.16)) показало, что доля ионной связи в ковалентной молекуле атома водорода равна 20%.

## 38.2 Пространственное строение молекул

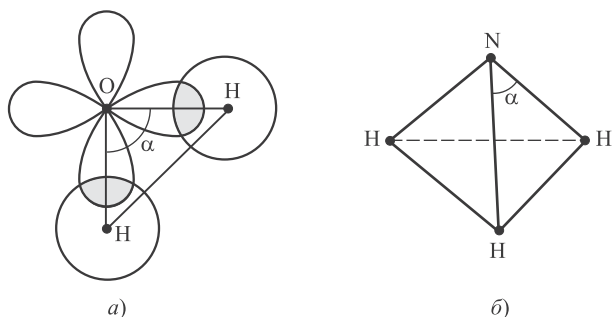
В молекуле водорода химическая связь двух атомов осуществляется, как мы видели, за счет влияния электронных облаков между атомами. В невозбужденных атомах эти электроны находятся в  $s$ -состояниях, и их волновые функции сферически симметричны. Форма перекрывания этих облаков показана на рис. 38.5, *a*.

Возможны случаи, когда валентный электрон находится в  $p$ -состоянии с угловым моментом, равным единице. Как следует из уравнения Шрёдингера, в этом состоянии волновая функция, соответствующая нулевой проекции момента на какую-либо ось, пропорциональна косинусу угла  $\vartheta$  между этой осью и радиусом-вектором. Следовательно, электронное облако имеет плотность, пропорциональную  $\cos^2 \vartheta$ . Если изобразить такое облако на рисунке, то оно будет иметь форму восьмерки (точнее, тела, образованного вращением восьмерки вокруг продольной оси). Если такой  $p$ -электрон спаривается с  $s$ -электроном другого атома, то последний стремится пристроиться таким образом, чтобы перекрытие волновых функций было максимальным. Очевидно, что такая связь осуществляется в направлении вытянутости облака  $p$ -электрона (рис. 38.5, *б*).

Электронное облако в состоянии  $m = 0$  «обволакивает» ось  $z$ , в двух других состояниях с проекциями  $m = \pm 1$  такие же облака вытянуты по осям  $x$  и  $y$ . Это помогает уяснить пространственное строение молекул различных веществ. Ниже мы приводим некоторые характерные примеры.



**Рис. 38.5.** Перекрывание электронных облаков (показано штриховкой) в молекуле водорода, когда оба электрона находятся в  $s$ -состоянии (*a*) и когда один из электронов находится в  $p$ -состоянии (*б*)



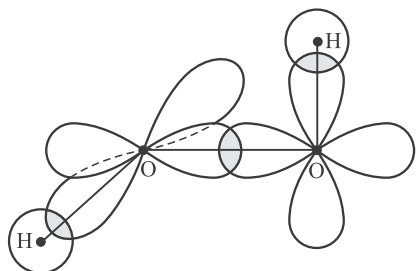
**Рис. 38.6.** Расположение атомов в молекуле воды (а) и в молекуле аммиака (б)

**Вода  $\text{H}_2\text{O}$ .** Атом кислорода имеет две заполненные оболочки  $1s$  и  $2s$ , на которых размещаются четыре электрона. Оставшиеся четыре валентных электрона находятся на оболочке  $2p$ . Их электронные облака отталкиваются и стремятся расположиться подальше друг от друга: три облака вытягиваются по осям  $x$ ,  $y$ ,  $z$  и четвертому деваться некуда. Скажем, оно располагается по оси  $z$ , причем электрон имеет противоположный спин по сравнению с другим электроном, облако которого также расположено вдоль этой оси. Образуются спаренные электроны, не участвующие в химической связи. Для химической связи остаются свободными два электрона — облака по осям  $x$  и  $y$ . Атомы водорода, подсоединяясь, как было сказано, образуют молекулу треугольной формы (рис. 38.6, а), причем угол  $\alpha$  между направлениями на атомы водорода должен быть равен  $90^\circ$ . Но частично «оголенные» протоны также отталкиваются, и этот угол несколько возрастает: его экспериментальное значение  $\alpha = 105^\circ$ . Это увеличение тем меньше, чем больше размер атома. Так, в аналогичном соединении  $\text{H}_2\text{S}$  атом серы больше атома кислорода, отталкивание протонов слабее и угол уменьшается до  $\alpha = 93^\circ$ . Еще больше атом селена, и в молекуле  $\text{H}_2\text{Se}$  атомы водорода расположены под углом  $\alpha = 90^\circ$ .

**Аммиак  $\text{NH}_3$ .** Атом азота имеет конфигурацию  $(1s)^2(2s)^2(2p)^3$ . Как и в кислороде, четыре электрона, находящиеся в  $1s$ - и в  $2s$ -состояниях, спарены и не участвуют в химической связи. Остаются три  $p$ -электрона, и все три связи расположены по осям  $x$ ,  $y$ ,  $z$ . При подсоединении трех атомов водорода образуется правильная треугольная пирамида с атомом азота в вершине (рис. 38.6, б). Из-за отталкивания протонов углы  $\alpha$  при вершине чуть больше  $90^\circ$ . У похожего соединения, фосфина  $\text{PH}_3$ , угол ближе к прямому, а у  $\text{AsH}_3$  просто  $\alpha = 90^\circ$ .

Из объемности молекулы аммиака следует существование двух вырожденных состояний, отличающихся положением атома азота по ту или иную сторону от основания пирамиды. Оба эти состояния соответствуют минимуму энергии. Следовательно, при перемещении атома азота на другую сторону от основания пирамиды энергия системы повышается. Иными словами, эти два состояния разделяет энергетический барьер. Образно говоря, при туннелировании сквозь этот барьер молекула аммиака словно выворачивается наизнанку, подобно зонтику под сильным порывом ветра.

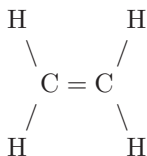
Квантово-механическое туннелирование снимает вырождение: низшие уровни энергии расщепляются, и переходы между ними приводят к излучению, которое сделало возможным создание аммиачного лазера.



**Рис. 38.7.** Форма и расположение электронных облаков в молекуле перекиси водорода

**Перекись водорода  $\text{H}_2\text{O}_2$ .** В случае, когда спариваются два  $p$ -электрона, максимальное перекрытие получается, если их «восьмерки» ориентированы вдоль одной оси. Так соединяются два атома кислорода в молекуле перекиси водорода (рис. 38.7). К ним подсоединяются обычным способом атомы водорода, причем вследствие отталкивания атомы водорода подсоединяются к кислородным облакам вдоль разных осей, так что линия  $\text{OO}$  и две линии  $\text{OH}$  на рис. 38.7 все взаимно ортогональны (на самом деле валентный угол между линиями  $\text{OO}$  и  $\text{OH}$  возрастает до  $100^\circ$ ).

В некоторых молекулах связи между атомами осуществляются не одной, а двумя или тремя парами электронов (двойные или тройные связи). Пример тройной связи — молекула  $\text{N}_2$ , которую химики записывают в виде  $\text{N}\equiv\text{N}$  (черточка соответствует паре электронов). Тройная связь атома углерода имеет место в молекуле ацетилена:  $\text{H}-\text{C}\equiv\text{C}-\text{H}$ . Двойная связь атома углерода осуществляется в молекуле этилена:



Квантовая механика объясняет и эти, и более сложные типы связей.

### 38.3 Молекулярные спектры

По сравнению с линейчатыми атомными спектрами молекулярные имеют более сложную структуру. Они состоят из совокупности полос, которые, в свою очередь, распадаются на ряд тесно расположенных линий. Причина такого усложнения спектров заключается в том, что наряду с движением электронов вокруг ядер в молекуле происходит колебание самих ядер относительно положения равновесия и вращение молекул как целого. Этим трем видам движений соответствуют три типа квантовых уровней энергии: электронные, колебательные и вращательные. В первом приближении движения можно считать независимыми. При изучении молекулярных спектров чрезвычайно важно, что масса ядра более чем на три порядка превышает электронную:  $M/m_e \sim 10^3 - 10^4$ . Поэтому скорости движения ядер в молекуле малы по сравнению со скоростями электронов. Движение электронов мгновенно приспособляется к изменению расположения

ядер. Таким образом, конфигурацию ядер можно рассматривать как параметр при определении уровней энергии  $E_{\text{эл}}$ . Полная энергия молекулы  $E$  складывается из *электронной, колебательной и вращательной* энергий, каждая из которых квантуется:

$$E = E_{\text{эл}} + E_{\text{кол}} + E_{\text{вр}}. \quad (38.17)$$

### Электронные уровни энергии

Молекула, подобно атому, обладает рядом возбужденных состояний, различающихся строением электронной оболочки. Переходы электрона из одного состояния в другое связаны с поглощением или испусканием квантов света. Порядок энергии электронных уровней молекулы можно оценить из соотношения неопределенностей Гейзенберга, аналогично тому, как была получена энергия основного состояния атома водорода. Если линейные размеры молекулы водорода  $R_0$ , то оценка дает

$$E_{\text{эл}} \sim \frac{(\Delta p)^2}{m_e} \approx \frac{\hbar^2}{m_e R_0^2}, \quad (38.18)$$

что при типичных значениях  $R_0 \sim 10^{-10}$  м составляет несколько электронвольтов и соответствует переходам с испусканием квантов в видимой и ультрафиолетовой областях спектра.

### Энергия колебательного движения ядер

Движение ядер в молекуле можно рассматривать при заданном электронном состоянии. Изучение относительного движения ядер с массами  $M_1$  и  $M_2$  в молекуле сводится к решению задачи о поведении частицы с приведенной массой  $M_{\text{пр}} = M_1 M_2 / (M_1 + M_2)$  в потенциальном поле. Для заданного связанного электронного состояния молекулы энергия  $E(R)$  имеет минимум, когда ядра расположены на расстоянии  $R_0$ . Разложим  $E(R)$  в ряд Тейлора вблизи точки минимума по степеням  $(R - R_0)$ :

$$E(R) = E(R_0) + \frac{1}{2} \left( \frac{\partial^2 E}{\partial R^2} \right)_{R_0} (R - R_0)^2 + \dots \quad (38.19)$$

С точностью до постоянной  $E(R_0)$  энергия движения ядер пропорциональна квадрату расстояния от положения равновесия. Следовательно, эти атомы колеблются в молекуле под действием квазиупругой силы

$$F = -\frac{\partial E}{\partial R} = -k_0(R - R_0). \quad (38.20)$$

Квантовая механика позволяет (см. разд. 36.5) определить энергетические состояния такой колебательной системы — гармонического осциллятора; этим состояниям соответствует совокупность эквидистантных (равноотстоящих друг от друга) уровней энергии

$$E_{\text{кол}} = \hbar\omega_0 \left( v + \frac{1}{2} \right), \quad v = 0, 1, 2, \dots, \quad (38.21)$$

где  $v$  — колебательное (вибрационное) квантовое число, а частота  $\omega_0 = \sqrt{k_0/M_{\text{пр}}}$  совпадает с частотой классического осциллятора. Переходы между колебательными уровнями подчиняются правилу отбора  $\Delta v = \pm 1$ , т. е. вибрационное число при переходе может изменяться лишь на единицу.

Для оценки порядка величины энергии колебательного движения заметим, что при амплитуде колебаний порядка расстояния  $R_0$  между ядрами молекула диссоциирует (разваливается). Энергия таких колебаний составляет величину порядка  $k_0 R_0^2$ . В то же время изменение энергии при амплитуде колебаний  $R_0$  близко к величине электронной энергии  $E_{\text{эл}}$ . Действительно, изменение межъядерного расстояния на величину  $R_0$  должно вызывать существенные искажения электронной волновой функции  $\Psi_+$ , т. е. возбуждение электронов с изменением их энергии на величину порядка  $E_{\text{эл}}$ . Таким образом, получаем оценку «коэффициента жесткости» молекулярного осциллятора:  $k_0 R_0^2 \sim E_{\text{эл}}$ , откуда  $k_0 \sim E_{\text{эл}}/R_0^2 \approx m_e E_{\text{эл}}^2/\hbar^2$  (см. соотношение (38.18)). Таким образом, для энергии колебательного движения получаем следующую оценку:

$$E_{\text{кол}} = \hbar\omega_0 = \hbar\sqrt{\frac{k_0}{M_{\text{пр}}}} \approx \sqrt{\frac{m_e}{M_{\text{пр}}}} E_{\text{эл}}, \quad (38.22)$$

что соответствует энергиям порядка десятых или сотых долей электронвольта и излучению в инфракрасной области спектра. Таким образом, энергии перехода между колебательными уровнями много меньше энергии перехода между электронными уровнями. Им соответствует обычно излучение в инфракрасном диапазоне.

### Энергия вращательного движения молекул

Вращение молекул рассмотрим в предположении жесткой связи между ядрами, т. е. пренебрегая колебаниями. Для двухатомной молекулы момент инерции относительно оси, перпендикулярной оси молекулы и проходящий через центр инерции, равен  $I = M_{\text{пр}} R_0^2/2$ . Согласно законам механики, энергия вращательного движения связана с моментом импульса  $L$  молекулы:  $E_{\text{вр}} = L^2/2I$ . Вращательный момент квантуется:  $L = \hbar\sqrt{J(J+1)}$ , где  $J = 0, 1, 2, \dots$  — ротационное квантовое число. Это позволяет определить вращательные уровни энергии:

$$E_{\text{вр}} = \frac{\hbar^2}{2I} J(J+1). \quad (38.23)$$

Отсюда следует оценка для вращательной энергии:

$$E_{\text{вр}} \sim \frac{\hbar^2}{2I} \approx \frac{\hbar^2}{M_{\text{пр}} R_0^2} \approx \frac{m_e}{M_{\text{пр}}} E_{\text{эл}}, \quad (38.24)$$

что дает величину порядка  $10^{-3}$ – $10^{-5}$  эВ, соответствующую излучению в далекой инфракрасной и микроволновой (СВЧ) областях спектра. Для вращательного спектра разрешены переходы с  $\Delta J = \pm 1$ .

## Молекулярные спектры

Проведенные оценки величины трех типов уровней в молекуле показывают, что

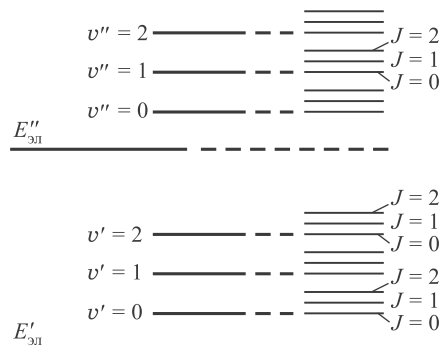
$$E_{\text{эл}} : E_{\text{кол}} : E_{\text{вр}} = 1 : \sqrt{\frac{m_e}{M_{\text{пр}}}} : \frac{m_e}{M_{\text{пр}}}. \quad (38.25)$$

В соответствии с этим система уровней молекулы складывается из сравнительно далеко отстоящих электронных уровней, испытывающих расщепление под влиянием колебаний ядер. Эти расщепленные уровни, в свою очередь, испытывают еще более тонкое расщепление из-за вращения молекул (рис. 38.8).

Изменение электронного состояния молекулы связано обычно с целой серией изменений: колебательное движение претерпевает изменения, так как новому состоянию соответствует измененное равновесное расстояние между ядрами; при этом также изменяется и момент инерции молекулы и, следовательно, положение вращательных уровней. Эти изменения вызывают целую серию поглощаемых или испускаемых квантов, энергии которых соответствуют разнице энергий каких-либо уровней, между которыми происходит переход.

Если атомные спектры состоят из отдельных линий, то при наблюдении с низким разрешением молекулярные спектры представляются состоящими из полос. При применении приборов с высокой разрешающей способностью обнаруживается, что полосы состоят из большого числа тесно расположенных линий. В соответствии с их характером спектры молекул носят название полосатых спектров. В зависимости от того, изменение каких видов энергии (электронной, колебательной или вращательной) вызывает испускание молекулой фотона, различают следующие типы полос: 1) вращательные; 2) колебательно-вращательные и 3) электронно-колебательные.

В основном состоянии молекулы все три вида энергии имеют минимальное значение. При сообщении молекуле достаточного количества энергии она переходит в возбужденное состояние и затем, совершая разрешенный правилами отбора переход в одно из более низких энергетических состояний, излучает фотон с энергией  $\hbar\omega_{\text{ф}} = \Delta E_{\text{эл}} + \Delta E_{\text{кол}} + \Delta E_{\text{вр}}$ . С учетом квантования трех видов энергии



**Рис. 38.8.** Схема расположения электронных, вращательных и колебательных уровней молекулы



мы можем записать это выражение в виде

$$\begin{aligned} \hbar\omega_{\Phi} = & E'_{\text{эл}} - E''_{\text{эл}} + \left(v' + \frac{1}{2}\right)\hbar\omega'_{\text{кол}} - \left(v'' + \frac{1}{2}\right)\hbar\omega''_{\text{кол}} + \\ & + \frac{\hbar^2 J'(J' + 1)}{2I'} - \frac{\hbar^2 J''(J'' + 1)}{2I''}. \end{aligned} \quad (38.26)$$

Следует отметить, что значения  $\omega_{\text{кол}}$  и  $I$  зависят от электронных конфигураций молекулы и поэтому обозначены одним и двумя штрихами.

Учитывая соотношения между энергиями  $\Delta E_{\text{эл}} \gg \Delta E_{\text{кол}} \gg \Delta E_{\text{вр}}$ , приходим к выводу, что при слабых возбуждениях изменяется только  $E_{\text{вр}}$ , при более сильных —  $E_{\text{кол}}$ , и лишь при еще более сильных возбуждениях изменяется электронная конфигурация молекулы, т. е.  $E_{\text{эл}}$ . Соответственно наименьшей энергией обладают фотоны, связанные с вращательными переходами (электронная конфигурация и энергия колебательного движения не изменяются). Учитывая, что  $\Delta J = \pm 1$ , находим:

$$\omega_{\Phi, \text{вр}} = \frac{\Delta E_{\text{вр}}}{\hbar} = \frac{\hbar}{2I} [(J+1)(J+2) - J(J+1)] = \frac{\hbar}{I}(J+1). \quad (38.27)$$

Измерение энергии таких фотонов позволяет определить момент инерции молекулы и, следовательно, ее размер  $R_0$ . Например, для HCl  $I = 2,71 \cdot 10^{-47}$  кг·м<sup>2</sup>, что соответствует  $R_0 = 1,29 \cdot 10^{-10}$  м = 1,29 Å.

При переходах, сопровождающихся изменениями и колебательного, и вращательного состояний молекулы, частота излучаемого фотона может быть записана в виде

$$\begin{aligned} \omega_{\Phi} = & \frac{\Delta E_{\text{кол}} + \Delta E_{\text{вр}}}{\hbar} = \omega_{\text{кол}} \left(v' + \frac{1}{2}\right) - \omega_{\text{кол}} \left(v'' + \frac{1}{2}\right) + \frac{\hbar}{I}(J+1) = \\ = & \omega_{\text{кол}} + \frac{\hbar}{I}(J+1), \end{aligned} \quad (38.28)$$

где было учтено, что  $\Delta v = \pm 1$ .

## 38.4 Комбинационное рассеяние света

**Комбинационное рассеяние** возникает при прохождении света через среду и состоит в том, что в спектре рассеяния симметрично несмещенной линии, соответствующей частоте падающего света  $\omega_0$ , появляются спутники — новые спектральные линии со смещенными частотами. Каждому красному спутнику с частотой  $\omega_0 - \omega_{\text{кол}}$  соответствует фиолетовый с частотой  $\omega_0 + \omega_{\text{кол}}$ . При обычных температурах интенсивность красных спутников на несколько порядков выше интенсивности фиолетовых, но с повышением температуры интенсивность фиолетовых спутников быстро возрастает.

В рамках классической электродинамики можно привести простые соображения о механизме возникновения спутников. Колебания молекулы описываются суперпозицией гармонических колебаний — нормальных мод (мы обсуждали их

в разделе 24.4, говоря о колебаниях связанных осцилляторов). Рассмотрим одну из мод с положением равновесия  $q_0$  и амплитудой  $q_1$ :

$$q(t) = q_0 + q_1 \cos \omega_{\text{кол}} t. \quad (38.29)$$

Поляризуемость  $\alpha$  молекулы (см. т. 1, гл. 17) зависит, вообще говоря, от ее колебаний и при небольших амплитудах может быть описана первыми членами разложения в ряд Тейлора:

$$\alpha \approx \alpha_0 + \alpha_1 \cos \omega_{\text{кол}} t, \quad \alpha_0 = \alpha(q_0), \quad \alpha_1 = \frac{\partial \alpha(q_0)}{\partial q_0} q_1. \quad (38.30)$$

Под действием электрического поля монохроматической волны  $\vec{E} = \vec{E}_0 \cos \omega_0 t$  молекула среды приобретает дипольный момент  $\vec{p} = \alpha \varepsilon_0 \vec{E}$  (см. т. 1, уравнение (17.20)). С учетом (38.30) это выражение можно записать в виде

$$\begin{aligned} \vec{p} &= \alpha \varepsilon_0 \vec{E} = (\alpha_0 + \alpha_1 \cos \omega_{\text{кол}} t) \varepsilon_0 \vec{E}_0 \cos \omega_0 t = \\ &= \varepsilon_0 \vec{E}_0 \left( \alpha_0 \cos \omega_0 t + \frac{\alpha_1}{2} \cos(\omega_0 - \omega_{\text{кол}}) t + \frac{\alpha_1}{2} \cos(\omega_0 + \omega_{\text{кол}}) t \right). \end{aligned} \quad (38.31)$$

Как мы уже знаем, колеблющийся диполь излучает на частоте своих колебаний. Поэтому первое слагаемое в (38.31) описывает упругое *рэлеевское рассеяние* света без изменения его частоты. Второе и третье слагаемые ведут к *комбинационному рассеянию света*, которое имеет *стоксову компоненту* с уменьшенной частотой и *антистоксову компоненту*, для которой частота больше.

С точки зрения квантовой теории рассеяние фотона с энергией  $\hbar \omega_0$  может сопровождаться переходами молекулы между различными вращательными или колебательными уровнями энергии. Если фотон с энергией  $\hbar \omega_0$  поглощается молекулой в основном состоянии, то она переходит в промежуточное виртуальное состояние<sup>1</sup>, откуда у нее есть лишь два пути: либо вернуться в основное состояние, испустив фотон той же частоты  $\omega_0$  (упругое рэлеевское рассеяние), либо попасть на колебательный уровень, лежащий выше основного на  $\hbar \omega_{\text{кол}}$ . В этом случае по закону сохранения энергии испускаемый «стоксов» фотон будет иметь меньшую частоту  $\hbar(\omega_0 - \omega_{\text{кол}})$ . Кроме того, начальный фотон может быть поглощен молекулой, уже находящейся на колебательном уровне с энергией  $\hbar \omega_{\text{кол}}$ . В этом случае молекула, возвращаясь из виртуального состояния в основное, испустит «антистоксов» фотон с увеличенной частотой  $\hbar(\omega_0 + \omega_{\text{кол}})$ . Так в спектре рассеяния излучения возникают серии симметрично расположенных красных и фиолетовых спутников.

Большая интенсивность красных спутников по сравнению с фиолетовыми объясняется тем, что в соответствии с распределением Больцмана при невысоких температурах число молекул, пребывающих в основном состоянии, намного превышает число молекул, находящихся в возбужденных состояниях. Поэтому

<sup>1</sup>Виртуальное состояние — ключевое понятие в квантовой теории поля: ненаблюдаемое промежуточное состояние системы, в котором нарушается обычная связь между энергией, импульсом и массой (см. гл. 45).

столкновения, сопровождающиеся увеличением энергии молекулы (уменьшением энергии фотона), происходят чаще, чем переходы в обратном направлении. При увеличении температуры число молекул в возбужденных колебательных состояниях растёт, что обуславливает увеличение интенсивности фиолетовых спутников.

Комбинационное рассеяние стало эффективным методом исследования строения молекул и их взаимодействия со средой. Методами комбинационного рассеяния изучают элементарные квантовые возбуждения в твердом теле — *квазичастицы* (подробнее см. разд. 40.5). Применение лазеров в качестве источников света значительно расширило круг объектов (газы, порошки), доступных для исследования методами комбинационного рассеяния света.

Комбинационное рассеяние света было открыто в 1928 г. Советские ученые Г. С. Ландсберг и Л. И. Мандельштам изучали прохождение света через прозрачные кристаллы, а индийские физики Ч. Раман и К. Кришнан — через газы и жидкости. В феврале 1928 г. обе группы практически одновременно (с разницей в неделю) обнаружили в спектре рассеяния красные и фиолетовые спутники. Несмотря на то что открытие было сделано независимо, Нобелевская премия была присуждена в 1930 г. одному Раману «за работы по рассеянию света и за открытие эффекта, названного в его честь». Так и повелось: в западной литературе этот эффект называется рамановским рассеянием, в отечественной — комбинационным рассеянием (а в немецкоязычной литературе встречается название «эффект Смекала—Рамана» по имени австрийского ученого А. Смекала, предсказавшего это явление в 1923 г.

## Часть VII

# Основы теории строения вещества

*Ныне зиждительных тел основных объясню я движенье,  
Коиm все вещи они порождают и вновь разлагают;  
Сила какая к тому принуждает их, скорость какая  
Свойственна им на пути в пустоте необъятной пространства,  
Ты же внимателен будь и выслушай то, что скажу я.*

---

Тит Лукреций Кар. «О природе вещей»



## Глава 39

# Взаимодействие излучения с веществом

В этой главе речь пойдет о физических принципах работы квантовых усилителей и генераторов. Принципиальная возможность существования квантового усилителя была отмечена в 1939 г. В. А. Фабрикантом, который обратил внимание на способность среды с инверсной населенностью уровней усиливать проходящее через нее излучение. Первые квантовые генераторы, работающие в диапазоне сантиметровых волн (*мазеры*), были созданы в 1953 г., а в 1960 г. был создан первый аналогичный прибор, работающий в оптическом диапазоне длин волн (*лазер*). Названия этих приборов происходят от первых букв английских слов: «*maser*» — *Microwave Amplification by Stimulated Emission of Radiation* (усиление микроволн с помощью вынужденного излучения), «*laser*» — *Light Amplification by Stimulated Emission of Radiation* (усиление света с помощью вынужденного излучения). В 1964 г. Н. Г. Басову, А. М. Прохорову (СССР) и Ч. Таунсу (США) была присуждена Нобелевская премия «за фундаментальные работы в области квантовой электроники, которые привели к созданию генераторов и усилителей на лазерно-мазерном принципе».

### 39.1 Вывод формулы Планка по Эйнштейну

Ранее мы обсудили, что атомы и ансамбли атомов могут находиться в различных состояниях, характеризуемых набором квантовых чисел. Переход из одного энергетического состояния  $E_n$  в другое  $E_m$  сопровождается поглощением или испусканием кванта электромагнитного излучения — фотона, обладающего энергией  $\hbar\omega = E_n - E_m$ , где  $\omega$  — частота излучения. Если  $E_n > E_m$ , то мы имеем дело со *спонтанным* (самопроизвольным) переходом атома с более высокого на более низкий уровень, в результате чего излучается фотон с энергией  $\hbar\omega$ . Схематически процесс *спонтанного излучения* можно изобразить в виде «реакции»  $A^* \rightarrow A + \gamma$ , где звездочка указывает на возбужденное состояние атома  $A$ .

Если же  $E_n < E_m$ , то мы имеем дело с **вынужденным переходом**, сопровождающимся **поглощением фотона**, энергия  $\hbar\omega$  которого идет на возбуждение атома (перевод его на более высокий энергетический уровень):  $\gamma + A \rightarrow A^*$ .

В 1918 г. А. Эйнштейн обратил внимание на то, что существуют «испускаемые» переходы другого типа, которые происходят под действием внешнего электромагнитного излучения. В таком процессе фотон падает на атом, находящийся в возбужденном состоянии, и заставляет его перейти в более низкое энергетическое состояние с излучением другого фотона. В конечном итоге в системе оказываются два фотона — начальный и излученный:  $\gamma + A^* \rightarrow A + \gamma + \gamma$ . Вероятность вынужденных переходов вниз возрастает с увеличением интенсивности стимулирующего излучения. Возникающее в результате таких переходов излучение называется **вынужденным (стимулированным, индуцированным) излучением**.

Спонтанное излучение хаотично по направлениям и фазам испускаемых фотонов, так как излучающие атомы независимы. Индуцированное же излучение должно быть по своим характеристикам совершенно тождественно с тем внешним излучением, которое, проходя через вещество, его породило. А именно индуцированное излучение имеет ту же частоту, направление и поляризацию, что и вынуждающее внешнее излучение. Фазы испускаемых фотонов скоррелированы с вынуждающими колебаниями, т. е. **вынужденное излучение когерентно**.

Выражения для вероятностей всех трех типов излучения рассчитываются в квантовой электродинамике. Во времена Эйнштейна эта теория еще не была создана, и он применил для анализа проблемы наглядные термодинамические соображения. Далее мы для простоты рассматриваем набор из  $N$  атомов, имеющих всего два уровня энергии  $E_1$  и  $E_2$  ( $E_2 > E_1$ ) — так называемую **двухуровневую среду**. Переход между уровнями сопровождается излучением или поглощением фотона.

Пусть в момент времени  $t$  какие-то  $N_2$  из общего числа атомов находятся в более высоком энергетическом состоянии 2 и пусть вероятность спонтанного перехода отдельного атома в более низкое (невозбужденное или основное) состояние в единицу времени равна  $A_{21}$ . Тогда изменение числа атомов в состоянии 2 за малое время  $dt$  составит  $dN_2 = -A_{21}N_2dt$ . Знак «минус» указывает на убыль числа атомов на уровне 2. Величина  $A_{21}$  называется **коэффициентом Эйнштейна** для спонтанного излучения. Теперь интегрированием легко получаем

$$N_2 = N_2(0)e^{-A_{21}t}, \quad (39.1)$$

где  $N_2(0)$  — число атомов в состоянии 2 в начальный момент времени. По смыслу формулы величина  $\tau = 1/A_{21}$  есть среднее время жизни атома в возбужденном состоянии (т. е. время, за которое число возбужденных атомов уменьшится в  $e$  раз). Этот параметр определяет характерное время затухания спонтанного излучения.

Представим теперь, что наши атомы находятся в равновесии с излучением частотой  $\omega = (E_2 - E_1)/\hbar$  и спектральной плотностью энергии  $u(\omega, T)$  (плотность энергии в единичном интервале частот). Спектральная плотность энергии пропорциональна числу фотонов данной частоты. Чем больше фотонов, тем вероятнее поглощение одного из них атомом. Поэтому для вероятности вынужденного

поглощения атомом излучения с частотой  $\omega$  в единицу времени можно написать выражение

$$W_{12} = B_{12}u(\omega, T), \quad (39.2)$$

где коэффициент Эйнштейна  $B_{12}$  характеризует свойства данного атома. Для числа переходов в возбужденное состояние за время  $dt$  имеем

$$dN_{12} = B_{12}u(\omega, T)N_1dt. \quad (39.3)$$

Для вероятности индуцированного излучения Эйнштейн предложил использовать формулу, аналогичную (39.2), поскольку, согласно его гипотезе, вынужденное излучение также пропорционально  $u(\omega, T)$ :

$$W_{21}^i = B_{21}u(\omega, T). \quad (39.4)$$

Коэффициент пропорциональности  $B_{21}$  может, вообще говоря, отличаться от  $B_{12}$ . Складывая  $W_{21}^i$  с вероятностью  $A_{21}$  спонтанного перехода, получаем полную вероятность перехода из состояния 2 в состояние 1 в единицу времени:

$$W_{21} = A_{21} + B_{21}u(\omega, T), \quad (39.5)$$

так что число переходов из возбужденного состояния в низшее за время  $dt$  равно:

$$dN_{21} = [A_{21} + B_{21}u(\omega, T)]N_2dt. \quad (39.6)$$

При термодинамическом равновесии вещества и электромагнитного поля должен соблюдаться баланс между процессами испускания и поглощения света, т. е. равенство полного числа актов испускания света и актов его поглощения. Такое равновесие устанавливается в замкнутой полости, температура  $T$  стенок которой поддерживается постоянной. Если в состоянии равновесия числа переходов  $2 \rightarrow 1$  и  $1 \rightarrow 2$  равны, т. е.

$$[A_{21} + B_{21}u(\omega, T)]N_2 = B_{12}u(\omega, T)N_1, \quad (39.7)$$

то мы получаем

$$u(\omega, T) = \frac{A_{21}}{B_{12} \frac{N_1}{N_2} - B_{21}}. \quad (39.8)$$

Числа атомов при термодинамическом равновесии подчиняются закону Больцмана (см. т. 1, разд. 13.4):

$$\frac{N_2}{N_1} = \exp\left(-\frac{E_2 - E_1}{k_B T}\right) = \exp\left(-\frac{\hbar\omega}{k_B T}\right). \quad (39.9)$$

Отсюда

$$u(\omega, T) = \frac{A_{21}}{B_{12} \exp\left(\frac{\hbar\omega}{k_B T}\right) - B_{21}}. \quad (39.10)$$



При повышении температуры спектральная плотность должна неограниченно возрастать. Так будет лишь при условии, если знаменатель в (39.10) стремится к нулю, а это возможно только при  $B_{21} = B_{12}$ , т. е. коэффициенты Эйнштейна для вынужденного поглощения и индуцированного излучения света равны<sup>1</sup>. Отсюда

$$u(\omega, T) = \frac{A_{21}/B_{21}}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1}. \quad (39.11)$$

Заметим также, что коэффициенты Эйнштейна не зависят от температуры, ибо относятся к отдельным актам поглощения — испускания фотонов атомом, а температура — это характеристика ансамбля атомов. В пределе высоких температур  $k_B T \gg \hbar\omega$  мы получаем из (39.11) выражение

$$u(\omega, T) \approx k_B T \frac{A_{21}}{\hbar\omega B_{21}}. \quad (39.12)$$

Из сопоставления предыдущей формулы с законом Рэлея—Джинса (33.25) следует тогда, что

$$\frac{A_{21}}{B_{21}} = \frac{\hbar\omega^3}{\pi^2 c^3}. \quad (39.13)$$

Подставляя (39.13) в (39.11), мы получаем формулу Планка (33.35) для спектральной плотности излучения черного тела. Соотношения между коэффициентами Эйнштейна, выведенные им из простых термодинамических соображений, были подтверждены впоследствии точными расчетами.

Разделив  $u(\omega, T)$  на число  $\omega^2/(\pi^2 c^3)$  типов колебаний в единице объема в единичном интервале частот (см. формулу (33.22)), получим *среднюю энергию одного типа колебания (фотонов с частотой  $\omega$ )*:

$$\langle \varepsilon_\omega \rangle = \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1}. \quad (39.14)$$

Разделив, в свою очередь, это выражение на энергию фотона  $\hbar\omega$ , найдем *среднее число фотонов данной частоты* при равновесии:

$$\langle n_\omega \rangle = \frac{1}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1}. \quad (39.15)$$

С этой формулой и ее аналогами мы еще встретимся в нашем курсе.

<sup>1</sup>Этот вывод справедлив для невырожденных уровней 1 и 2. Если это не так и кратности вырождения уровней 1 и 2 равны  $g_1$  и  $g_2$  соответственно, то модифицируется формула Больцмана (39.9), а с ней и соотношение между коэффициентами Эйнштейна:  $g_1 B_{12} = g_2 B_{21}$ .

## 39.2 Общие сведения о лазерах

### Прохождение излучения через вещество.

#### Инверсия населенностей

Снова рассмотрим двухуровневую среду с энергетическими уровнями  $E_2$  и  $E_1$  ( $E_2 > E_1$ ). Если на эту среду падает монохроматическое излучение частотой  $\omega = (E_2 - E_1)/\hbar$ , то при распространении его на расстояние  $dx$  изменение интенсивности (потока энергии) излучения будет связано с поглощением излучения атомами и вынужденным излучением атомов. За счет вынужденного испускания интенсивность проходящего через среду излучения возрастает, так как растет число фотонов. Прирост интенсивности  $dI_+$  должен быть пропорционален начальной интенсивности, концентрации  $n_2$  атомов на верхнем уровне энергии, пройденному расстоянию  $dx$  и коэффициенту Эйнштейна  $B_{21}$ :  $dI_+ = \beta I n_2 B_{21} dx$ . Здесь  $\beta$  — размерный коэффициент пропорциональности. Аналогично за счет процессов поглощения фотонов плотность энергии в пучке уменьшается пропорционально концентрации атомов на нижнем уровне и коэффициенту Эйнштейна  $B_{12}$ :  $dI_- = -\beta I n_1 B_{12} dx$ . Складывая  $dI_+$  и  $dI_-$ , находим полное изменение  $dI$  интенсивности проходящего через среду излучения:

$$dI = \beta I (B_{21} n_2 - B_{12} n_1) dx. \quad (39.16)$$

Учитывая равенство коэффициентов Эйнштейна  $B_{21} = B_{12}$  и вводя **коэффициент поглощения**  $\alpha$ , записываем это уравнение в виде

$$dI = -\alpha I dx, \quad \alpha = \beta B_{21} (n_1 - n_2). \quad (39.17)$$

Решение дифференциального уравнения (39.17) имеет вид

$$I = I_0 e^{-\alpha x}. \quad (39.18)$$

Эта формула дает изменение плотности энергии  $I$  в потоке фотонов при прохождении ими слоя вещества толщиной  $x$  (значение  $I_0$  соответствует точке  $x = 0$ ).

В термодинамическом равновесии в соответствии с распределением Больцмана,  $n_2 < n_1$  ( $E_2 > E_1$ ), и поэтому коэффициент поглощения  $\alpha$  всегда положителен:

$$\alpha = \beta B_{21} n_1 \left( 1 - \frac{n_2}{n_1} \right) = \beta B_{21} n_1 \left( 1 - e^{-\frac{E_2 - E_1}{k_B T}} \right) > 0. \quad (39.19)$$

Таким образом, плотность энергии излучения, как видно из (39.19), убывает по мере прохождения через вещество, т. е. свет поглощается. Однако если создать систему, в которой  $n_2 > n_1$ , то коэффициент поглощения  $\alpha$  станет отрицательным и будет иметь место не ослабление, а *усиление интенсивности* света. Состояние системы, в котором  $n_2 > n_1$ , называется **состоянием с инверсной населенностью уровней**, а вещество в таком состоянии называется **активной средой**. Инверсная населенность уровней противоречит равновесному распределению Больцмана и может быть создана искусственно, если система выведена из состояния термодинамического равновесия. Это создает принципиальную возможность

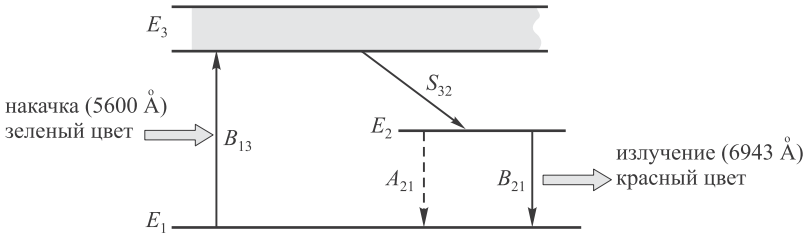


Рис. 39.1. Схема активной трехуровневой среды (рубин)

усиления и генерации когерентного оптического излучения и используется на практике при разработке источников такого излучения — лазеров.

## Принцип работы лазера

Создание лазера стало возможным после того, как были найдены способы осуществления инверсной населенности уровней в некоторых веществах (активных средах). Первый рабочий оптический квантовый генератор (лазер) был создан в США Т. Мейманом (1960) на основе рубина. Рубин представляет собой кристаллическую решетку  $\text{Al}_2\text{O}_3$ , содержащую небольшую (0,03% — 0,05%) примесь ионов хрома ( $\text{Cr}^{3+}$ ). На рисунке 39.1 представлена схема энергетических уровней хрома (*трехуровневая среда*). Широкий уровень  $E_3$  используется для возбуждения ионов хрома светом мощной газоразрядной лампы с широкой полосой частот в зелено-голубой области видимого света — *лампы накачки*. Возбуждение ионов хрома  $\text{Cr}^{3+}$  за счет энергии накачки от внешнего источника изображено стрелкой  $B_{13}$ .

Электроны с короткоживущего уровня  $E_3$  совершают быстрый ( $10^{-7}$  —  $10^{-8}$  с) безызлучательный переход на уровень  $E_2$  (изображен стрелкой  $S_{32}$ ). Выделяемая при этом энергия не испускается в виде фотонов, а передается кристаллу рубина. При этом рубин нагревается, поэтому в конструкции лазера предусматривается охлаждение (рис. 39.2). Время жизни долгоживущего узкого уровня  $E_2$  составляет примерно  $10^{-3}$  с, т. е. на пять порядков больше, чем у широкополосного уровня  $E_3$ . При достаточной мощности накачки число электронов на уровне  $E_2$  (его называют *метастабильным уровнем*) становится больше, чем на уровне  $E_1$ , т. е. создается инверсная населенность между «рабочими» уровнями  $E_2$  и  $E_1$ .

Фотон, излученный при спонтанном переходе между этими уровнями (изображен стрелкой  $A_{21}$ ), индуцирует испускание дополнительных (вынужденных) фотонов (переход показан стрелкой  $B_{21}$ ), которые, в свою очередь, вызывают вынужденное излучение целого каскада фотонов с длиной волны  $\lambda = 6943 \text{ \AA}$ .

**Задача 39.59.** Определить относительную населенность  $n_2/n_1$  рабочих уровней в кристалле рубина при комнатной температуре в условиях термодинамического равновесия.

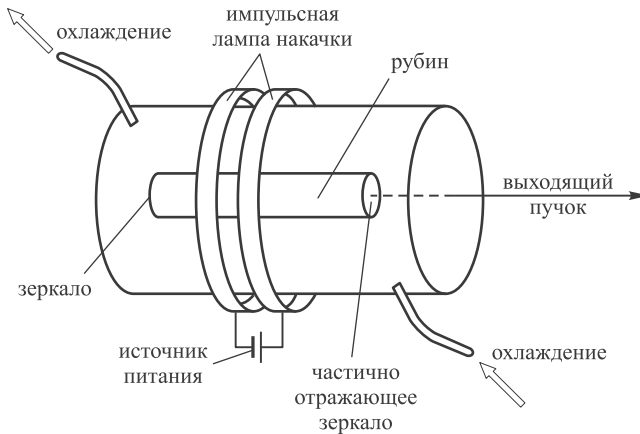


Рис. 39.2. Схема рубинового лазера

**Решение.** Исходя из длины волны, испускаемой рубиновым лазером, находим разность энергий рабочих уровней:

$$\Delta E = E_2 - E_1 = h\nu = \frac{hc}{\lambda} = \frac{6,63 \cdot 10^{-34} \times 3 \cdot 10^8}{6,943 \cdot 10^{-7}} = 2,86 \cdot 10^{-19} \text{ Дж.}$$

При комнатной температуре  $T = 300 \text{ К}$  имеем:  $k_B T = 4,14 \cdot 10^{-21} \text{ Дж}$ . Из распределения Больцмана следует:

$$\frac{n_2}{n_1} = e^{-\Delta E / (k_B T)} = e^{-69} = 10^{-30}. \blacksquare$$

Однако создание активной среды с инверсной населенностью уровней — это лишь половина дела. Необходимо также создать режим генерации света, т. е. использовать *положительную обратную связь*. Активная среда сама по себе способна лишь усиливать проходящее излучение. Для осуществления режима генерации необходимо такое усиление индуцированного излучения, которое компенсировало бы все потери в системе. Для этого активное вещество помещают в *оптический резонатор*, образованный, как правило, двумя параллельными зеркалами, одно из которых является полупрозрачным и служит для вывода излучения из резонатора. Другое зеркало, имеющее высокий коэффициент отражения, возвращает излучение обратно в активную среду, где оно продолжает усиливаться: так работает положительная обратная связь. Конструктивно в первых лазерах на рубине использовались кристаллы цилиндрической формы длиной примерно 40 мм и диаметром около 5 мм. Торцы были отполированы параллельно друг другу и служили зеркалами резонатора. Один из торцов был посеребрен так, что коэффициент отражения был близок к единице, а другой торец был полупрозрачным (т. е. имел коэффициент отражения заметно меньше единицы) — через него лазерное излучение выходило наружу.

Источником возбуждения служила мощная импульсная ксеноновая лампа накачки, обвивающая рубин спиралью (см. рис. 39.2). При достаточной мощности лампы больше половины ионов хрома переводится в возбужденное состояние.

После того, как достигается инверсная населенность рабочих уровней с энергиями  $E_2$  и  $E_1$ , первые спонтанно излучаемые фотоны, соответствующие переходу между этими уровнями, не имеют преимущественного направления распространения. Они вызывают вынужденное излучение, также распространяющееся по всем направлениям в кристалле рубина (напомним, что фотоны возникающие при вынужденном излучении, летят в том же направлении, что и падающие фотоны).

Фотоны, направления движения которых образуют малые углы с осью оптического резонатора (в данном случае с осью рубинового стержня), испытывают многократные отражения от его торцов и создают вынужденное излучение. Фотоны же, распространяющиеся в других направлениях, выходят из кристалла рубина через его боковую поверхность и не участвуют в формировании выходного излучения. Так в резонаторе генерируется узкий луч света, а многократное прохождение фотонов через активную среду индуцирует испускание все новых и новых фотонов, усиливая интенсивность выходного излучения.

Таким образом, оптический резонатор выполняет две функции: во-первых, создает положительную обратную связь и, во-вторых, формирует узкий направленный пучок излучения с определенной пространственной структурой.

В рассмотренной трехуровневой схеме для создания инверсной населенности между рабочими уровнями нужно возбудить больше половины атомов, что требует значительных затрат энергии. Более эффективной является *четырёхуровневая схема*, которая также имеет место в твердотельных лазерах — например, использующих кристаллы, активированные ионами неодима  $\text{Nd}^{3+}$ . В наиболее распространенном газовом лазере — *гелий-неоновом* — также выполняются условия для генерации излучения по четырехуровневой схеме. Активной средой в таком лазере является смесь инертных газов — гелия и неона — с энергией основного состояния  $E_1$  (которую мы принимаем за точку отсчета энергий, полагая  $E_1 = 0$ ). Накачка осуществляется в процессе электрического газового разряда, в результате которого атомы гелия переходят в возбужденное состояние с энергией  $E_3 = 20,61$  эВ. Лежащий чуть выше уровень  $E_4 = 20,66$  эВ в атоме неона случайно оказался близок к уровню  $E_3$  в гелии (рис. 39.3). При столкновении атомов гелия с атомами неона недостающая энергия возбуждения может быть эффективно передана последним без излучения. Таким образом, уровень  $E_4$  неона оказывается более населенным, чем нижележащий уровень  $E_2 = 18,70$  эВ. Переход между ними сопровождается излучением с длиной волны  $632,8$  нм (красная линия в спектре), которая является основной в промышленных He—Ne-лазерах. На уровне  $E_2$  атомы неона долго не задерживаются, быстро возвращаясь в основное состояние. Заметим, что уровень  $E_2$  в неоне населен крайне незначительно, и потому для создания инверсной населенности между  $E_4$  и  $E_2$  надо возбудить небольшое число атомов гелия. Это требует гораздо меньших затрат энергии как на накачку, так и на охлаждение установки, что является характерной особенностью лазеров, работающих по четырехуровневой схеме генерации. Для работы лазера могут быть использованы и другие уровни неона (на рис. 39.3 не показаны), дающие излучение как в видимом, так и в инфракрасном (ИК) диапазоне, причем гелий используется только для процесса накачки.

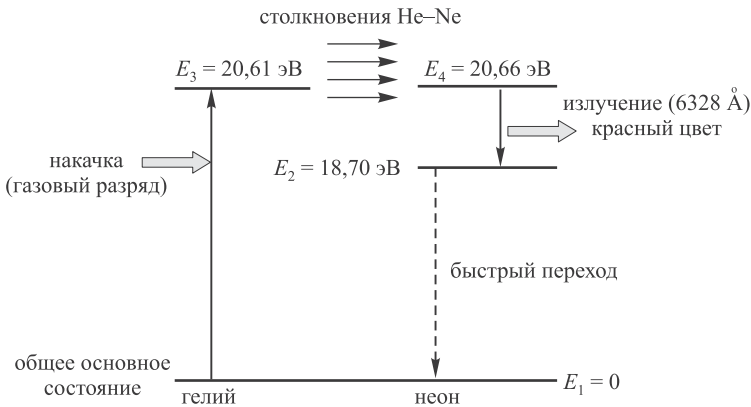


Рис. 39.3. Схема четырехуровневого He–Ne-лазера

**Задача 39.60.** Найти относительную населенность уровня  $E_2$  в неоне при комнатной температуре.

**Решение.** Эта задача отличается от предыдущей лишь численными значениями. Для разнообразия проведем вычисления в электронвольтах. Выразим сначала в этих единицах постоянную Больцмана:  $k_B = 1,3806 \cdot 10^{-23} / 1,602 \cdot 10^{-19} = 8,617 \cdot 10^{-5}$  эВ/К, так что при комнатной температуре  $k_B T = 25,85$  мэВ. Теперь легко находим:

$$\frac{n_2}{n_1} = e^{-18,70/25,85 \cdot 10^{-3}} = e^{-723,4} \sim 10^{-314}.$$

Столь малое число с практической точки зрения не отличается от нуля, и даже при легкой накачке создается инверсная населенность уровня  $E_4$ . ■

Излучение лазеров отличается характерными особенностями: высокой временной и пространственной когерентностью (монохроматичность излучения и малая расходимость пучка) и высокой спектральной интенсивностью. Характеристики излучения зависят от типа лазера и режима работы, однако можно отметить некоторые близкие к предельным значения параметров:

- ширина линии генерации  $\sim 0,1$  Å;
- пиковая интенсивность  $\sim 10^{13}$  Вт/мм<sup>2</sup>;
- короткие импульсы до  $\sim 10^{-16}$  с;
- расходимость пучка ( $\sim 1''$ ).

Короткие (пикосекундные) импульсы лазера незаменимы при изучении быстротекущих процессов. В импульсе может развиваться большая мощность (до нескольких гигаватт), равная мощности нескольких блоков АЭС по миллиону киловатт каждый. При этом излучение может быть сосредоточено в узком конусе углов. Такие пучки с большой частотой повторения импульсов незаменимы в промышленности при изготовлении с высокой точностью деталей из труднообрабатываемых (особо твердых или хрупких) материалов.

## Типы лазеров

В рамках курса общей физики мы не можем остановиться подробно на специфических особенностях, бытовых, технических и научных применениях лазеров

различных типов ввиду их чрезвычайного многообразия. Ограничимся лишь кратким обзором некоторых типов лазеров, различающихся характеристиками активной среды и способами накачки. Важно понимать, что, несмотря на многообразие типов лазеров и их конструктивные особенности, необходимыми элементами любого лазера являются: активная среда, т. е. среда, в которой может быть создана инверсная населенность уровней; система накачки, с помощью которой инверсная населенность создается; оптический резонатор, обеспечивающий необходимую для процесса генерации положительную обратную связь и одновременно формирующий структуру лазерного пучка.

**Твердотельные лазеры.** Обычно они бывают импульсными, первым таким лазером был описанный выше рубиновый. Популярны лазеры на стекле с неодимом в качестве рабочего вещества. Они генерируют свет с длиной волны  $\sim 1,06$  мкм, имеют большие размеры и пиковую мощность порядка тераватт. Примеры — огромные лазеры в военных лабораториях США: «Шива» и «Нова» в Ливерморе, «Трайидент» в Лос-Аламосе, использовавшиеся для экспериментов по управляемому термоядерному синтезу.

Распространены лазеры на кристаллах итрий-алюминиевого граната, излучающие в ИК-диапазоне на длине волны 1,064 мкм. Они могут работать как в непрерывном режиме генерации, так и в импульсном, с частотой повторения импульсов до нескольких килогерц (для сравнения: рубиновый лазер производит один импульс в несколько минут). В качестве активаторов лазерных кристаллов и стекол используются ионы редкоземельных элементов или группы железа. Такие лазеры имеют широкий спектр применений в электронной технике (лазерная технология), оптической локации, медицине и др.

**Газовые лазеры.** Обычно это лазеры непрерывного действия. Отличаются правильной пространственной структурой пучка. Пример: гелий-неоновый лазер, генерирующий свет на длинах волн 0,63, 1,15 и 3,39 мкм и имеющий мощность порядка милливатт. В технике широкой используется  $\text{CO}_2$ -лазер с мощностью порядка киловатт и выше и длинами волн 9,6 и 10,6 мкм. Газовые лазеры характеризуются наименьшей расходимостью пучка излучения. Один из способов накачки газовых лазеров — электрический разряд. Газовые лазеры обычно работают в непрерывном режиме, однако так называемые щелевые  $\text{CO}_2$ -лазеры с накачкой высокочастотным разрядом обеспечивают режим работы с частотой повторения импульсов до 20 кГц при средней мощности 500 Вт и высокой пиковой мощности. Такие лазеры с успехом используются, в частности, для резки металлов. Разновидность лазеров с активной газовой средой — химические и эксимерные лазеры.

**Химические лазеры.** Инверсная населенность создается в процессе химической реакции между двумя газами, например водородом (дейтерием) и фтором. В основе лежат экзотермические реакции  $\text{H} + \text{F}_2 \rightarrow \text{HF} + \text{F}$  и  $\text{F} + \text{H}_2 \rightarrow \text{HF} + \text{H}$ . Молекулы HF уже образуются в колебательно-возбужденном состоянии, что сразу создает инверсную населенность. Образовавшаяся рабочая смесь пропускается со сверхзвуковой скоростью через оптический резонатор, в котором в виде электромагнитного излучения выделяется часть накопленной энергии. С помощью системы зеркал резонатора это излучение фокусируется в узкий луч. Такие лазеры излучают большую энергию (более 2 кДж), длительность импульса у них порядка 30 нс, мощность до 200 ГВт. КПД (химический) достигает 10%, в то время как обычный КПД других лазеров — доли

процента. Генерируемая длина волны — 2,8 мкм (3,8 мкм для лазеров на фтористом дейтерии DF).

Из многочисленных проектов химических лазеров наиболее перспективными признаны лазеры на фтористом водороде либо на фтористом дейтерии. Проблемы: излучение лазеров на фтористом водороде с длиной волны 2,8 мкм активно рассеивается молекулами воды, всегда содержащимися в атмосфере. Это намного ослабляет яркость излучения. Лазер на фтористом дейтерии работает на длине волны, для которой атмосфера практически прозрачна. Однако удельное энерговыделение лазеров на DF в 1,5 раза меньше, чем у лазеров на HF. Это значит, что при использовании их в космосе придется доставлять на орбиту гораздо большее количество химического топлива.

**Эксимерные лазеры.** Молекулы эксимера — это двухатомные молекулы (например,  $\text{Xe}_2$ ), которые могут находиться только в возбужденном состоянии (невозбужденное состояние оказывается у них неустойчивым). С этим связана основная особенность эксимерных лазеров: основное состояние эксимерных молекул является незаполненным, т. е. нижний рабочий лазерный уровень всегда оказывается пустым. Накачка осуществляется импульсным электронным пучком: он переводит значительную часть атомов в возбужденное состояние, в котором атомы объединяются в эксимерные молекулы. Появление молекул в возбужденном состоянии автоматически создает инверсную населенность уровней.

Поскольку переход между рабочими уровнями является широкополосным, возможна перестройка частоты генерации. Лазер на  $\text{Xe}_2$  дает перестраиваемое излучение в области УФ ( $\lambda = 172$  нм) и имеет высокую эффективность ( $\sim 20\%$ ) преобразования энергии. В настоящее время эксимерные ArF-лазеры с длиной волны  $\lambda = 193$  нм используются в офтальмологической хирургии для поверхностного испарения (абляции) роговицы.

**Жидкостные лазеры.** Активное вещество в жидком состоянии однородно и допускает циркуляцию с целью охлаждения, что создает преимущества перед твердотельными лазерами. Это позволяет получить большие энергии и мощности в импульсном и непрерывном режимах. В первых жидкостных лазерах (1964–1965) использовались соединения редкоземельных элементов. На смену им пришли лазеры на растворах органических красителей.

В таких лазерах обычно используется оптическая накачка излучением других лазеров, работающих в видимом или УФ-диапазоне. Интересным свойством лазеров на красителях является возможность перестройки частоты генерации в широких пределах. Подбором красителя можно получить излучение на любой длине волны от ближнего ИК- до ближнего УФ-диапазона. Это связано с широкими сплошными колебательно-вращательными спектрами молекул жидкости. Особенностью таких лазеров являются сверхкороткие (менее пикосекунды) импульсы. В этой области они когда-то произвели революцию, но сейчас уступают место новейшим технологиям — твердотельным лазерам на кристаллах титана-сапфира и волоконным лазерам на фотонных кристаллах.

**Полупроводниковые лазеры.** В отдельный класс выделяются твердотельные лазеры на полупроводниковых материалах. Накачка производится бомбардировкой пучком электронов, мощным лазерным излучением, но чаще методами электроники. В полупроводниковых лазерах используются переходы не между дискретными уровнями энергии, а между разрешенными энергетическими зонами, т. е. совокупностями близко расположенных уровней (об энергетических зонах в кристаллах подробнее говорится в последующих разделах.) Использование различных полупроводниковых материалов позволяет получить излучение на длинах волн от 0,7 до 1,6 мкм. Размеры активной среды чрезвычайно малы: длина резонатора может быть менее 1 мм (хорошо известный пример полупроводникового лазера — лазерная указка).



Типичная мощность порядка нескольких киловатт (у самых распространенных красных лазерных указок мощность гораздо меньше — от 1 до 20 мВт), эффективность достигает 50%. Имеют широкое применение (волоконная оптика, связь), могут быть использованы для проецирования телевизионных изображений на большой экран.

**Лазеры на свободных электронах.** Пучок высокоэнергетических электронов пропускается через «магнитную гребенку», вынуждающую электроны колебаться с заданной частотой. Соответствующее устройство — ондулятор — представляет собой ряд магнитов, которые располагаются между секциями ускорителя, так что релятивистские электроны движутся вдоль оси ондулятора (от фр. *onde* — волна) и совершают поперечные ей колебания, излучая первичную («спонтанную») электромагнитную волну. В открытом резонаторе, куда далее поступают электроны, спонтанная электромагнитная волна усиливается, создавая когерентное лазерное излучение. Главная особенность лазеров на свободных электронах состоит в возможности плавной перестройки частоты генерации (от видимого до ИК-диапазона) за счет изменения кинетической энергии электронов. КПД таких лазеров составляет примерно 1% при средней мощности до 4 Вт. С использованием устройств возврата электронов в резонатор КПД может быть увеличен до 20–40%. Лазеры на свободных электронах родственны лампе бегущей волны, клистрону и прочим генераторам СВЧ.

**Рентгеновский лазер с ядерной накачкой.** Это наиболее экзотический лазер. Схематично он представляет собой ядерную боеголовку, на поверхности которой укреплено до 50 металлических стержней, ориентированных в разных направлениях. Стержни имеют две степени свободы и, подобно стволам орудий, могут направляться в любую точку пространства. Вдоль оси каждого стержня расположена тонкая проволока из материала высокой плотности (порядка плотности золота) — активная среда. Источником энергии накачки лазера служит ядерный взрыв. При взрыве активное вещество переходит в плазменное состояние. Мгновенно остывая, плазма испускает когерентное излучение в мягком рентгеновском диапазоне. Из-за высокой концентрации энергии излучение, попадая на цель, приводит к взрывному испарению вещества, образованию ударной волны и разрушению корпуса цели.

В описанном лазере не предусматриваются зеркала резонатора, использование которых в рентгеновском диапазоне не представляется возможным. Это устройство одноразового использования, принцип действия и конструкция которого делают очевидными сферы его применения. Во время подземных ядерных взрывов в 1980-е гг. на полигоне в штате Невада (США) специалисты Национальной лаборатории им. Лоуренса (Ливермор) провели оценочные испытания первых рентгеновских лазеров с ядерной накачкой. Википедия, если ей можно верить, приводит следующие параметры лазерного излучения: длина волны около 14 Å, длительность импульса  $10^{-9}$  с, мощность излучения рентгеновского лазера при атомном взрыве превышает 400 ТВт (<http://tinyurl.com/zvnjxpv>). Этот проект, получивший кодовое название «Экスカлибур» по названию легендарного меча не менее легендарного короля Артура, был закрыт с окончанием «холодной войны». Так что пока, к счастью, потенциально мощное новое оружие скорее выглядит как фантазия по мотивам фильма «Звездные войны».

### 39.3 Нелинейные эффекты в оптике

Распространение в среде световых волн, излучаемых обычными источниками света, описывается линейными дифференциальными уравнениями (линейная оптика), что означает независимость оптических характеристик среды от интенсивности света. Создание лазеров сделало возможным получение световых волн с

напряженностью электрического поля, сравнимой по величине с напряженностью микроскопического внутриатомного поля ( $\sim 10^{10}$  В/м). В таких полях показатель преломления и другие оптические характеристики среды обнаруживают зависимость от напряженности электрического поля  $E$  световой волны. В этом случае нарушается принцип суперпозиции полей, т. е. принцип независимого распространения электромагнитных волн в веществе, а соответствующие дифференциальные уравнения становятся **нелинейными**. Это приводит к существенному изменению характера протекания в среде известных оптических явлений, а также к возникновению совершенно новых, так называемых **нелинейных эффектов в оптике**. К ним относятся:

- нелинейное поглощение света (самопросветление и самозатемнение среды);
- преобразование частоты (генерация высших гармоник, параметрическая генерация и вынужденное рассеяние света);
- эффекты самовоздействия (самофокусировка и самодефокусировка пучков);
- многофотонные процессы.

Нелинейный эффект **насыщения поглощения света** наблюдался задолго до появления лазеров. В 1923 г. С. И. Вавилов и В. Л. Левшин обнаружили уменьшение коэффициента поглощения света урановым стеклом с ростом интенсивности облучения. Этот первый нелинейный эффект в оптике объясняется тем, что вследствие сильного облучения значительная доля поглощающих частиц среды переходит в возбужденное состояние и уже не может далее поглощать свет.

Нелинейные эффекты **преобразования частоты** и **самовоздействия света** возникают из-за того, что в поляризации среды появляется нелинейная составляющая, которая растет с увеличением интенсивности электромагнитной волны. Наличие такой нелинейной составляющей объясняется ангармонизмом колебаний частиц среды в поле мощной световой волны. Из-за этого ангармонизма отклик среды на гармоническое внешнее электрическое поле становится нелинейным, т. е. перестает повторять форму внешнего воздействия. Таким образом, в электрическом поле, переизлученном атомами и молекулами вещества, появляются новые гармонические составляющие — происходит **генерация высших оптических гармоник** падающего на среду оптического излучения. Это легко понять, если представить, что в разложении электрической индукции  $D$  по степеням напряженности электрического поля  $E$  в сильных полях уже недостаточно ограничиваться лишь слагаемым первого порядка (линейным членом), но надо учесть и следующие члены ряда:

$$D = \varepsilon_0 (\varepsilon E + \xi E^2 + \zeta E^3 + \dots). \quad (39.20)$$

Подставляя сюда выражение для монохроматической волны  $E = E_0 \cos \omega t$  и учитывая формулы тригонометрии

$$\cos^2 \alpha = \frac{1 + \cos 2\alpha}{2}, \quad \cos^3 \alpha = \frac{3 \cos \alpha + \cos 3\alpha}{4}$$

и т. п., мы приходим к выводу, что в нелинейной среде появляются световые волны с частотами, кратными частоте падающей волны:  $\omega$ ,  $2\omega$ ,  $3\omega$  и т. д. Наибольшее распространение в практических приложениях имеет процесс **генерации второй гармоники** в «нелинейных» кристаллах со специально подобранными

свойствами. Так, например, не видимое глазом ИК-излучение YAG:Nd<sup>3+</sup>-лазера с длиной волны  $\lambda = 1,064$  мкм в кристалле KDP (дигидрофосфат калия) преобразуется в зеленый свет с длиной волны  $\lambda/2 = 0,532$  мкм.

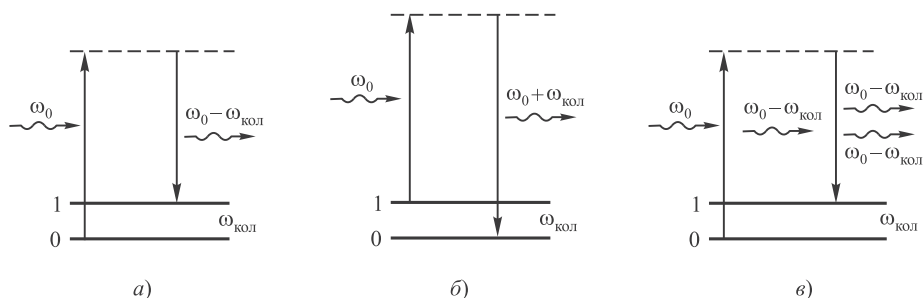
Нелинейная поляризация среды является также причиной зависимости показателя преломления от интенсивности электромагнитной волны. Если показатель преломления  $n$  увеличивается с ростом интенсивности, то лучи изгибаются в сторону оси пучка — туда, где  $n$  больше (см. задачу 28.14 в разделе 28.2). При превышении некоторой критической интенсивности наблюдается явление *самофокусировки света* — расходящийся световой пучок становится сходящимся. Явление самофокусировки было предсказано в 1961 г. советским физиком-теоретиком Г. А. Аскарьяном и в 1965 г. впервые наблюдалось Н. Ф. Пилипецким и А. Р. Рустамовым. Возможен и противоположный процесс *самодефокусировки* пучка, если показатель преломления уменьшается с ростом интенсивности волны.

При *многофотонных процессах* нелинейного взаимодействия оптического излучения с веществом в одном элементарном акте происходит поглощение (или испускание) двух и более фотонов. Вероятность таких процессов увеличивается с ростом интенсивности волны. *Многофотонный фотоэффект* или *многофотонная фотоионизация* сопровождаются поглощением нескольких фотонов, энергия каждого из которых меньше работы выхода или энергии ионизации соответственно.

К нелинейным процессам взаимодействия со средой относится и явление *вынужденного комбинационного рассеяния света*. В разделе 38.4 мы обсуждали комбинационное рассеяние света, которое происходит самопроизвольно и называется *спонтанным комбинационным рассеянием*. Поэтому рассеянное излучение некогерентно и очень слабо. В таком линейном приближении учитывается движение молекул под влиянием только электрического поля внешнего излучения. Но в сильных полях лазерного излучения на молекулу действует результирующее поле  $\vec{E}$ , которое складывается из поля  $\vec{E}_0 \cos \omega_0 t$  внешнего излучения и поля  $\vec{E}_{0\text{ст}} \cos(\omega_0 - \omega_{\text{кол}})t$  стоксова рассеянного излучения, которым уже нельзя пренебречь. Дипольный момент молекулы сам зависит от поля  $\vec{E}$ , а сила, действующая на диполь, пропорциональна произведению дипольного момента на градиент поля. Таким образом, в уравнениях движения молекулы появляется член, пропорциональный квадрату результирующего поля  $\vec{E}$ :

$$\vec{E}^2 = \left( \vec{E}_0 \cos \omega_0 t + \vec{E}_{0\text{ст}} \cos(\omega_0 - \omega_{\text{кол}})t \right)^2. \quad (39.21)$$

Если раскрыть это выражение, то получим суперпозицию колебаний с разными частотами. Квадраты косинусов ведут к удвоению частоты рассеянного излучения — этот эффект мы уже разобрали выше. Произведение же косинусов ведет к появлению колебаний с частотами, равными сумме и разности частот внешнего поля  $\omega_0$  и стоксова излучения  $\omega_0 - \omega_{\text{кол}}$ . Суммарная частота  $\omega_0 + (\omega_0 - \omega_{\text{кол}}) \approx 2\omega_0$  — это снова генерация второй гармоники. А вот разность частот  $\omega_0 - (\omega_0 - \omega_{\text{кол}})$  в точности равна частоте  $\omega_{\text{кол}}$  собственных колебаний молекулы. Таким образом, в сильных полях обратное действие стоксова излучения ведет к резонансной раскачке колебаний молекул и соответственно к росту интенсивности рассеянного излучения. Кроме того, в условиях резонанса



**Рис. 39.4.** Схема комбинационного рассеяния: стоксова (а), антистоксова (б) и вынужденного (в)

вынужденные колебания в молекуле происходят в фазе с падающей волной, так что возникающее вынужденное комбинационное излучение когерентно с внешним излучением. Его интенсивность сравнима с интенсивностью внешнего излучения, в отличие от спонтанного комбинационного рассеяния, интенсивность которого на несколько порядков меньше интенсивности накачки.

С точки зрения квантовой теории процессы комбинационного рассеяния представлены на рис. 39.4, где цифрами 0 и 1 показаны основной и возбужденный колебательные уровни с частотой перехода между ними  $\omega_{\text{кол}}$ , а штриховой линией — виртуальные уровни. При спонтанном стоксовом рассеянии (рис. 39.4, а) фотон лазерного излучения частотой  $\omega_0$  переводит молекулу с уровня 0 на некоторый виртуальный энергетический уровень, откуда она возвращается на возбужденный уровень 1. При этом испускается фотон меньшей частоты  $\omega_0 - \omega_{\text{кол}}$ . При спонтанном антистоксовом рассеянии (рис. 39.4, б) лазерный фотон переводит на виртуальный (более высокий) уровень молекулу, уже находившуюся в возбужденном состоянии 1. При возвращении молекулы с виртуального уровня на основной излучается фотон большей частоты  $\omega_0 + \omega_{\text{кол}}$ . Наконец, при вынужденном комбинационном рассеянии молекула, переведенная на виртуальный (нижний) уровень лазерным фотоном, возвращается на возбужденный уровень под действием уже испущенного ранее стоксова фотона, в результате чего в системе оказываются уже два фотона с одинаковой частотой  $\omega_0 - \omega_{\text{кол}}$ . Если такая среда помещена в оптический резонатор, то процесс вынужденного комбинационного рассеяния (рис. 39.4, в) нарастает лавинообразно, пока не будет достигнуто насыщение.

Имеются и другие нелинейные процессы, происходящие при взаимодействии мощного лазерного излучения со средой. Их перечисление можно продолжать долго, но мы завершим эту главу констатацией факта, что с появлением лазеров возникла новая область физики — **нелинейная оптика**.

## Контрольные вопросы

1. Чем различаются спонтанное и вынужденное излучения?
2. Почему для работы лазера необходимо создание инверсной населенности уровней?

3. Почему время жизни широкополосного уровня  $E_3$  в рубиновом лазере намного меньше времени жизни узкополосного уровня  $E_2$ ?
4. В чем роль уровня  $E_3$  в рубиновом лазере, если он не относится к числу «рабочих», как уровни  $E_2$  и  $E_1$ ?
5. Излучение лазера представляет собой почти параллельный пучок лучей. Падает ли интенсивность излучения по закону обратного квадрата расстояния?
6. Чем излучение лазера лучше пучка света от сильной лампы, сфокусированного обычной линзой?
7. Каковы преимущества четырехуровневой схемы работы лазера?
8. Почему нелинейные оптические эффекты, как правило, не наблюдаются с обычными световыми источниками?
9. Поясните на квантовом языке процессы спонтанного и вынужденного комбинационного рассеяния.
10. Почему можно не учитывать антистоксову компоненту при рассмотрении вынужденного комбинационного рассеяния света?

## Глава 40

# Теплоемкость кристаллов. Статистика Бозе—Эйнштейна

Классическая статистика, основанная на распределении Максвелла—Больцмана, подсчитывала число частиц, находящихся в малом объеме  $dV$  вокруг некой точки  $\vec{r}$  и имеющих импульсы в малом интервале  $d\vec{p}$  вокруг некоего значения  $\vec{p}$ . Познакомившись с законами квантовой механики, мы теперь понимаем, что для квантовых систем такой подход неприменим: микрочастицы не могут одновременно иметь определенные значения координаты и импульса, так что необходимо учесть соотношение неопределенностей Гейзенберга. Кроме того, мы уже знакомы с принципом запрета Паули, согласно которому электроны «мешают» друг другу и не могут занять одно и то же состояние. Очевидно, что это также должно найти отражение в квантовой статистике, знакомству с которой посвящена эта глава. Первое и почти очевидное применение квантовой статистики нашла в теории теплоемкости твердых тел. Мы коснемся здесь также интереснейших макроскопических квантовых явлений — сверхтекучести и сверхпроводимости.

### 40.1 Классические представления о теплоемкости кристаллов

Простейшей моделью кристалла является геометрически правильно построенная кристаллическая решетка, в узлах которой помещаются атомы, считающиеся материальными точками. Атомы совершают тепловые колебания около положений равновесия. Если колебания малы, то их можно считать гармоническими. Энергия каждого атома складывается из кинетической и потенциальной энергий. На каждую степень свободы приходится в среднем кинетическая энергия  $k_B T/2$  и такое же количество средней потенциальной энергии. Таким образом, среднее значение полной энергии, приходящейся на одну колебательную степень свободы, равно (см. также т. 1, разд. 11.3):

$$\langle E \rangle = \langle E_{\text{к}} \rangle + \langle E_{\text{пот}} \rangle = k_B T.$$

Напомним теперь классические результаты для теплоемкости кристаллической решетки, о которых говорилось в разделе 11.3. Предположим для простоты, что все атомы одинаковы и каждый из них обладает тремя колебательными степенями свободы, а потому на него приходится средняя энергия  $3k_B T$ . Умножив эту величину на постоянную Авогадро  $N_A$ , можно получить внутреннюю энергию теплового движения молекул одного моля твердого тела:

$$U_m = N_A \times 3k_B T = 3RT, \quad (40.1)$$

где  $R$  — универсальная газовая постоянная. Отсюда для молярной теплоемкости твердого тела имеем

$$c_m = \frac{\partial U_m}{\partial T} = 3R. \quad (40.2)$$

Этот закон был установлен (1819) французскими учеными П. Дюлонгом и А. Пти как эмпирическое правило, согласно которому произведение удельной теплоемкости химического элемента в твердом состоянии на его молярную массу примерно одинаково для всех элементов и равно  $3R = 25$  Дж/(моль · К).

Если вещество кристалла состоит из молекул, имеющих  $n_a$  атомов, то **закон Дюлонга и Пти** (40.2) модифицируется очевидным образом:  $c_m = 3n_a R$ . Так, для молекул поваренной соли NaCl  $n_a = 2$  и  $c_m \approx 50$  Дж/(моль · К), что близко к реальному значению  $c_m \approx 51,4$  Дж/(моль · К) при  $20^\circ$ .

### Недостатки классической теории теплоемкости

- Классическая теория не дает объяснения зависимости теплоемкости тел от температуры. Экспериментально установлено, что при приближении к абсолютному нулю теплоемкости всех тел, в том числе и кристаллов, стремятся к нулю.
- По теореме о равномерном распределении энергии все степени свободы равноправны. Однако опытные факты приводят к заключению, что при низких температурах вклад в теплоемкость вносят не все из них: при понижении температуры определенные степени свободы становятся малоэффективными (говорят, что они «заморожены»). Так, при температуре  $T \leq T_{\text{кол}} = \hbar\omega/k_B$  «замерзают» колебательные степени свободы (здесь  $\omega$  — характерная частота колебаний). Аналогично при дальнейшем понижении температуры «замерзают» вращательные степени свободы, это происходит при  $T \leq T_{\text{вращ}} = \hbar^2/(Ik_B)$ , где  $I$  — момент инерции молекулы.
- Закон Дюлонга и Пти был сформулирован для любого твердого тела: металла и диэлектрика. Однако металл состоит из положительно заряженных ионов, совершающих тепловые колебания вблизи узлов кристаллической решетки. Между ними движутся так называемые свободные электроны, которые ведут себя подобно электронному газу. Наличием свободных электронов объясняется высокая электрическая проводимость металлов. Классическая теория теплоемкости не принимает во внимание наличие электронного газа в металлах. Она учитывает тепловые колебания только ионов и приводит к результату Дюлонга и Пти. В отличие от диэлектриков,

в металлах, казалось бы, следует учитывать вклад в теплоемкость, вносимый электронами. На каждый свободный электрон приходится средняя кинетическая энергия  $3k_B T/2$ . Поэтому по классической теории теплоемкость электронного газа должна быть сравнимой с теплоемкостью кристаллической решетки. Однако опыт показывает, что свободные электроны практически не вносят никакого вклада в теплоемкость металлов.

Более подробно эти вопросы будут рассмотрены ниже.

## 40.2 Средняя энергия квантового осциллятора

В этом разделе мы разберемся с «замораживанием» колебательных степеней свободы осциллятора. Рассмотрим систему из множества независимых квантовых осцилляторов с одной частотой  $\omega$ . В квантовой механике было показано, что энергия осциллятора квантуется, т. е. принимает ряд дискретных значений:

$$E_n = \hbar\omega \left( n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad (40.3)$$

В соответствии с общими принципами статистической физики вероятность  $P_n$  найти осциллятор в состоянии, характеризуемом неким значением  $n$  колебательного квантового числа, определяется формулой Больцмана

$$P_n = C \exp \left( -\frac{E_n}{k_B T} \right) = C \exp \left[ -\frac{\hbar\omega}{k_B T} (n + 1/2) \right], \quad (40.4)$$

где  $C$  — нормировочная постоянная. Для ее определения надо воспользоваться условием нормировки вероятности  $\sum_{n=0}^{\infty} P_n = 1$ . В итоге имеем

$$P_n = \left[ 1 - e^{-\hbar\omega/(k_B T)} \right] e^{-n\hbar\omega/(k_B T)}. \quad (40.5)$$

Если полное число осцилляторов в нашей системе равно  $N$ , то среднее число  $N_n$  осцилляторов, находящихся в состоянии с главным квантовым числом  $n$ , равно:

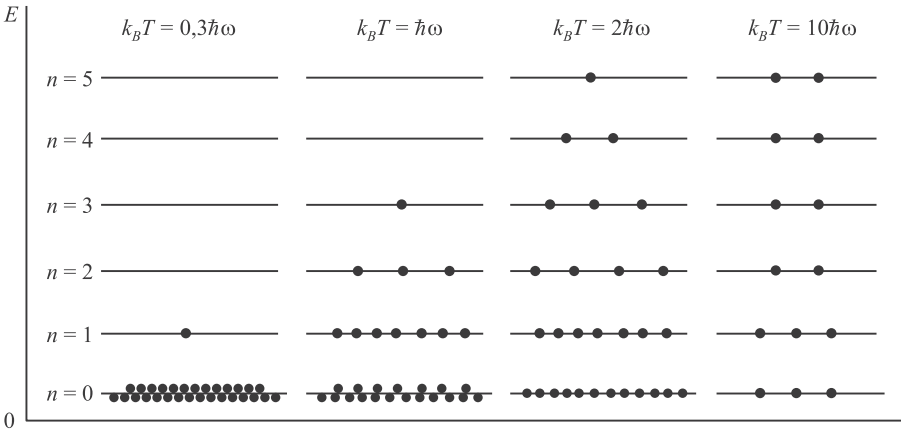
$$N_n = N P_n = N \left[ 1 - e^{-\hbar\omega/(k_B T)} \right] e^{-n\hbar\omega/(k_B T)}. \quad (40.6)$$

Видно, что чем больше значение квантового числа  $n$ , тем меньше вероятность обнаружить осциллятор в таком состоянии. Чем выше температура, тем большие значения  $n$  становятся практически значимыми для системы. При  $T \rightarrow 0$  к нулю стремятся все вероятности  $P_n$  с  $n \geq 1$  и лишь  $P_0 \rightarrow 1$ . Иными словами, при нулевой температуре нет тепловых возбуждений, и осциллятор находится в основном состоянии с наименьшей энергией  $E_0 = \hbar\omega/2$ .

Распределение осцилляторов по энергиям в зависимости от температуры системы показано на рис. 40.1. Для наглядности мы взяли систему из небольшого ( $N = 30$ ) числа осцилляторов (строго говоря, статистические законы применимы к системам с гораздо большим числом частиц).

После этого остается найти среднюю энергию нашей системы осцилляторов:  $\langle E \rangle = \sum_{n=0}^{\infty} E_n P_n$ . Такие суммы мы уже вычисляли (разд. 33.4), когда говорили





**Рис. 40.1.** Примерное распределение  $N = 30$  квантовых осцилляторов по энергетическим уровням в зависимости от температуры  $T$ . Показаны только основной и пять возбужденных уровней энергии. При  $T = 0$  все осцилляторы находятся в основном состоянии. При температуре, равной  $T = 0, 3\hbar\omega/k_B$ , наиболее вероятно найти только один осциллятор в первом возбужденном состоянии, а остальные, скорее всего, по-прежнему будут находиться в основном состоянии. По мере роста температуры становятся доступными более высокие энергии и распределение осцилляторов по уровням становится все более равномерным

о формуле Планка для излучения абсолютно черного тела. Здесь получаются идентичные результаты за исключением появления энергии нулевых колебаний. Сначала найдем *среднее значение*  $\langle n \rangle$  колебательного квантового числа  $n$  при некоторой температуре  $T$ :

$$\langle n \rangle = \sum_{n=0}^{\infty} n P_n = \frac{1}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1}. \quad (40.7)$$

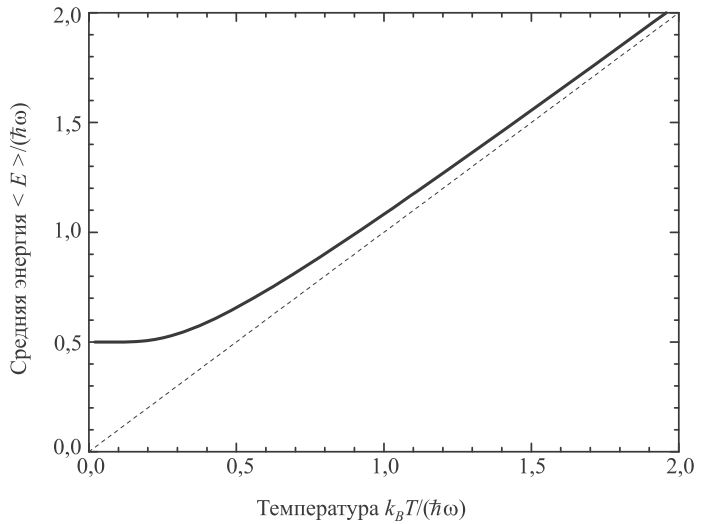
Теперь легко получить среднюю энергию осциллятора:

$$\langle E \rangle = \hbar\omega \left( \langle n \rangle + \frac{1}{2} \right) = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1} = \frac{\hbar\omega}{2} \operatorname{cth} \left( \frac{\hbar\omega}{2k_B T} \right). \quad (40.8)$$

Функция *гиперболический котангенс* определяется соотношением

$$\operatorname{cth} x = \frac{e^x + e^{-x}}{e^x - e^{-x}}. \quad (40.9)$$

На рисунке 40.2 изображена средняя энергия квантового осциллятора в единицах  $\hbar\omega$  как функция «безразмерной температуры»  $\theta = k_B T / \hbar\omega = T / T_{\text{кол}}$  (сплошная линия). Пунктирная линия  $\langle E \rangle = k_B T$  соответствует результату классической физики.



**Рис. 40.2.** Средняя энергия квантового осциллятора в зависимости от температуры

Видно, что квантовые поправки важны при низких температурах: при  $\theta < 0,3$  средняя энергия осциллятора близка к энергии основного состояния  $\hbar\omega/2$ . В таком случае говорят, что колебательные степени свободы «заморожены», т. е. тепловой энергии недостаточно для возбуждения колебаний. Но уже при  $\theta \simeq 2$  обе энергии практически совпадают, т. е. квантовые поправки малы. Значение  $\theta = 1$  можно принять за условную границу между квантовой и классическими областями. Ее смысл прозрачен: при  $k_B T = \hbar\omega$  тепловая энергия равна минимальной энергии возбуждения осциллятора, т. е. разности между энергией  $E_1 = 3\hbar\omega/2$  первого возбужденного состояния и энергией  $E_0 = \hbar\omega/2$  основного состояния осциллятора.

Повторим еще раз: «замораживание» колебательных степеней свободы происходит из-за существования энергетической «щели» между основным состоянием и первым возбужденным уровнем. В классической механике возможны сколь угодно малые изменения энергии и эффект «замораживания» степеней свободы отсутствует.

Какие же температуры можно считать низкими для осциллятора, моделирующего реальную систему, например молекулу водорода  $\text{H}_2$ ? Характерные частоты молекулярных колебаний располагаются обычно в инфракрасной области и имеют порядок  $\nu \approx 10^{14}$  Гц. Этому соответствуют энергия  $\hbar\omega = h\nu \approx 6 \cdot 10^{-20}$  Дж = = 0,4 эВ и температура  $T_{\text{кол}} \approx \hbar\omega/k_B \approx 4000$  К. Очевидно, что это весьма высокая температура, и поэтому в обычных условиях мы эффекта «замораживания» колебательных степеней свободы не наблюдаем.

## Средняя энергия квантового ротатора

Таким образом, привычные для нас комнатные температуры оказываются слишком низкими для возбуждения колебаний молекул. Посмотрим теперь, что происходит с молекулами при температурах  $T < T_{\text{кол}}$ . Так как колебания отсутствуют,

двухатомную молекулу можно представить в виде «гантели» — двух атомов, жестко соединенных между собой (см. разд. 38.3). Такая система называется *ротатором* и, как мы видели в разделе 10.4, имеет пять степеней свободы — три поступательных (движение центра масс) и две вращательных. Энергия вращательного движения классического ротатора имеет вид  $E_{\text{вр}} = I\omega^2/2$ , где  $I$  — момент инерции молекулы (см. т. 1, разд. 5.3). Учитывая связь  $I\omega = L$  между угловой частотой вращения  $\omega$ , моментом инерции  $I$  и моментом количества движения  $L$ , записываем классическую энергию вращения молекулы как  $E_{\text{вр}} = L^2/(2I)$ . В квантовой механике квадрат момента количества движения квантуется:  $L = \hbar\sqrt{J(J+1)}$ , где  $J = 0, 1, 2, \dots$  — *ротационное квантовое число*. Поэтому квантуется и энергия вращательного движения молекулы:

$$E_{\text{вр}} = \frac{\hbar^2}{2I}J(J+1). \quad (40.10)$$

Используя это соотношение и распределение Больцмана, можно получить выражение для средней энергии квантового ротатора. Однако в этом случае формулы достаточно сложны и мы ограничимся качественными результатами.

Как и в случае квантового осциллятора, здесь также имеется «щель» между основным и первым возбужденным состояниями. При высоких температурах средняя энергия стремится к классическому значению  $k_B T$ , соответствующему двум степеням свободы (вращение вокруг двух ортогональных осей). При низких температурах ротатор будет находиться в основном состоянии, соответствующем значению  $J = 0$  (отсутствие вращения). Переход между двумя этими предельными случаями осуществляется, очевидно, при такой температуре  $T_{\text{вр}}$ , когда тепловое движение способно возбудить вращательное движение. Минимальная (отличная от нуля) энергия вращения равна  $E_1 = \hbar^2/I$ , как это следует из (40.10) при  $J = 1$ . Поэтому  $T_{\text{вр}} \sim E_1/k_B = \hbar^2/Ik_B$ . Для момента инерции молекулы можно принять оценку  $I \approx 2m_p a_B^2$ , где  $m_p = 1,67 \cdot 10^{-27}$  кг (масса протона), а  $a_B = 5 \cdot 10^{-11}$  м (радиус Бора). Получаем тогда  $I \approx 8,4 \cdot 10^{-48}$  кг·м<sup>2</sup>,  $E_1 \approx 1,3 \cdot 10^{-21}$  Дж и  $T_{\text{вр}} \approx 100$  К.

Полученные оценки подтверждаются измерениями молярной теплоемкости при постоянном объеме  $c_{m,V}$ , которые мы уже обсуждали в т. 1, разд. 11.3. При температурах ниже 100 К в тепловом движении участвуют только поступательные степени свободы молекулы. В этом случае средняя энергия молекулы равна  $3k_B T/2$ , а энергия одного моля —  $3N_A k_B T/2 = 3RT/2$ , откуда следует выражение для теплоемкости  $c_{m,V} = 3R/2$ . В диапазоне температур 100–200 К молярная теплоемкость увеличивается до значения  $c_{m,V} = 5R/2$ , что свидетельствует о «размораживании» двух дополнительных (вращательных) степеней свободы (т. е. о добавлении энергии  $k_B T$  на молекулу). При температурах 4000–5000 К молярная теплоемкость снова увеличивается, на этот раз до значения  $c_{m,V} = 7R/2$ . Это «разморозилась» колебательная степень свободы, что принесло дополнительную энергию  $k_B T$  на молекулу.

### 40.3 Теплоемкость кристаллов по Эйнштейну

Выражение (40.8) было положено А. Эйнштейном в основу квантовой теории теплоемкости твердых тел. Эйнштейн отождествил кристаллическую решетку из  $N$  молекул с системой  $3n_a N$  независимых гармонических осцилляторов с одинаковой собственной частотой  $\omega$ . Тогда внутренняя энергия одного моля определяется выражением  $U = 3n_a N_A \langle E \rangle$ . Дифференцируя по температуре, получаем молярную теплоемкость кристаллической решетки твердых тел:

$$c_m = \frac{\partial U}{\partial T} = 3n_a R \left( \frac{\hbar\omega}{k_B T} \right)^2 \frac{e^{\hbar\omega/(k_B T)}}{[e^{\hbar\omega/(k_B T)} - 1]^2} = 3n_a R \left[ \frac{\frac{\hbar\omega}{2k_B T}}{\operatorname{sh} \left( \frac{\hbar\omega}{2k_B T} \right)} \right]^2, \quad (40.11)$$

где функция  $\operatorname{sh}x = (e^x - e^{-x})/2$  — гиперболический синус. Это и есть **формула Эйнштейна для теплоемкости кристаллов**.

При высоких температурах, когда  $\hbar\omega \ll k_B T$ , она переходит в классическую формулу  $c_m = 3n_a R$ . В другом предельном случае — при низких температурах  $\hbar\omega \gg k_B T$  — можно пренебречь убывающей экспонентой в гиперболическом синусе (т. е. считать  $\operatorname{sh}x \approx e^x/2$ ) и получить

$$c_m \approx 3n_a R \left( \frac{\hbar\omega}{k_B T} \right)^2 \exp \left( -\frac{\hbar\omega}{k_B T} \right). \quad (40.12)$$

При  $T \rightarrow 0$  полученное выражение стремится к нулю, как это требует тепловая теорема Нернста.

Поясним физический смысл этого результата. Как уже говорилось, из-за квантовой дискретности между основным и возбужденным уровнями системы осцилляторов имеется энергетическая «щель»  $\hbar\omega$ . Меньшее количество энергии осциллятор воспринять просто не в состоянии. При нулевой температуре в системе нет возбуждений — все осцилляторы находятся в основном состоянии. При небольшом повышении температуры тепловой энергии не хватает на преодоление этой щели, и лишь малое количество осцилляторов, пропорциональное  $\exp(-\hbar\omega/(k_B T))$  согласно закону Больцмана, переходит на первый возбужденный уровень. Именно они ответственны за поглощение тепловой энергии и наличие низкой теплоемкости кристалла. При высоких температурах тепловой энергии хватает на возбуждение многих вышележащих уровней, так что дискретность энергии уже не играет особой роли — мы возвращаемся к классическому результату Дюлонга и Пти.

Однако согласие с опытом теории Эйнштейна имеет только качественный характер. В выражении для  $c_m$  при низких температурах экспоненциальный множитель убывает быстрее, чем растет множитель  $1/T^2$ . Поэтому при приближении температуры к абсолютному нулю величина  $c_m$  стремится к нулю практически по экспоненциальному закону. Опыт же показывает, что теплоемкость кристаллов изменяется при низких температурах не экспоненциально, а по закону  $c_m \propto T^3$ . Как оказалось, эти расхождения теории Эйнштейна с опытом связаны не с существом квантовой теории, а с упрощением расчета, в котором предполагалось, что все гармонические осцилляторы колеблются с одной и той же частотой.

На самом деле кристаллическую решетку следует рассматривать как связанную систему взаимодействующих частиц. При вычислении теплоемкости тело действительно можно рассматривать как систему гармонических осцилляторов, но с *различными частотами*. Тем самым задача сводится к отысканию спектра частот.

#### 40.4 Теория теплоемкости твердых тел Дебая

Дебай предположил, что колебания атомов в кристаллической решетке не являются независимыми. Смещение одного из атомов из положения равновесия влечет за собой смещение других соседних с ним атомов. Таким образом, кристалл представляет собой систему  $N$  упруго связанных между собой атомов, обладающих  $3N$  степенями свободы. Каждая степень свободы (нормальное колебание) может быть представлена как гармонический осциллятор, среднюю энергию которого  $\langle E \rangle$  мы уже вычислили. Из-за связи между атомами частоты нормальных колебаний уже не совпадают между собой. Взаимодействие атомов приводит к тому, что колебание, возникшее в каком-то месте кристалла, передается от одного атома к другому, в результате чего возникает упругая волна. Эта волна, дойдя до границы кристалла, отражается. При наложении прямой и отраженной волн образуется стоячая волна, которой соответствует нормальное колебание кристаллической решетки. Число  $dN$  нормальных колебаний, т. е. стоячих волн, в интервале частот от  $\omega$  до  $\omega + d\omega$  велико, поэтому суммирование в выражении для внутренней энергии системы может быть заменено интегрированием:

$$U = \int \langle E \rangle dN. \quad (40.13)$$

#### Число колебаний в единице объема

В этом разделе мы займемся подсчетом числа стоячих волн, имеющих близкие частоты  $\omega$ . В сущности, мы выполнили уже эти выкладки для электромагнитного излучения (см. разд. 33.3), но повторим их снова с небольшими модификациями для применения также и к звуковым колебаниям в кристалле.

Рассмотрим сначала одномерный потенциальный ящик длиной  $L_x$ . Мы могли уже убедиться, что стоячая волна в нем (неважно, электромагнитная ли, звуковая или волна де Бройля) описывается функцией  $\sin(kx)$ , которая должна обращаться в нуль на границах ящика. Отсюда

$$k_x = \frac{\pi n}{L_x}. \quad (40.14)$$

Число  $n_x$  нумерует различные стоячие волны вдоль оси  $x$ , и потому на малый интервал волнового вектора  $dk_x$  приходится число колебаний

$$dn_x = \frac{L_x dk_x}{2\pi}. \quad (40.15)$$

Двойку в знаменателе мы поставили, чтобы избежать двойного счета: замена  $k_x$  на  $-k_x$  приводит к той же стоячей волне. В трехмерном ящике для волн, распространяющихся по другим осям, получаем аналогичные формулы

$$dn_y = \frac{L_y dk_y}{2\pi}, \quad dn_z = \frac{L_z dk_z}{2\pi}. \quad (40.16)$$

Перемножая (40.15) и (40.16), находим для полного числа стоячих волн в ящике объемом  $V = L_x L_y L_z$ :

$$dN = dn_x dn_y dn_z = \frac{V}{(2\pi)^3} d^3 k. \quad (40.17)$$

Наконец, учтем, что каждой стоячей волне может соответствовать  $g$  поляризаций (для частиц со спином  $s$  имеем  $g = 2s + 1$  — число различных проекций спина). Окончательно имеем

$$dN = \frac{gV}{(2\pi)^3} d^3 k. \quad (40.18)$$

Формула (40.18) дает число различных стоячих волн (отличающихся числом узлов и направлениями поляризации) в объеме  $V$ , приходящихся на элемент объема  $d^3 k$  пространства значений волнового вектора. Далее для перехода к частотам волн вспомним соотношение  $k = \omega/v$ , где  $v$  — фазовая скорость волны. Отсюда  $dk = d\omega/v$  и

$$dN = \frac{gV}{(2\pi)^3} 4\pi k^2 dk = \frac{gV}{2\pi^2 v^3} \omega^2 d\omega. \quad (40.19)$$

Мы вывели формулу (40.19) для прямоугольного объема, но можно показать, что форма объема не влияет на результат. Не имеет большого значения и физическая природа колебаний, число которых мы подсчитали. Например, для фотонов  $v = c$  и  $g = 2$  (свет может иметь правую и левую циркулярные поляризации). В итоге получаем уже известную нам формулу (33.22) для числа типов фотонов в объеме  $V$  и в интервале частот  $d\omega$ :

$$dN = \frac{V}{\pi^2 c^3} \omega^2 d\omega. \quad (40.20)$$

Для применения (40.19) к звуковым волнам в кристалле учтем, что там возможна одна продольная волна, распространяющаяся со скоростью  $v_{\parallel}$ , и две поперечные волны с разными поляризациями, как у фотонов, распространяющиеся со скоростью  $v_{\perp}$ . Очевидно теперь, как обобщить формулу (40.19) на данный случай:

$$\begin{aligned} dN &= \frac{2V}{2\pi^2 v_{\perp}^3} \omega^2 d\omega + \frac{V}{2\pi^2 v_{\parallel}^3} \omega^2 d\omega = \frac{V}{2\pi^2} \left( \frac{2}{v_{\perp}^3} + \frac{1}{v_{\parallel}^3} \right) \omega^2 d\omega = \\ &= \boxed{V \frac{3}{2\pi^2 v^3} \omega^2 d\omega}. \end{aligned} \quad (40.21)$$

Здесь мы ввели величину  $v$ , играющую роль некой средней величины между скоростями продольных и поперечных волн; она вычисляется из соотношения

$$\boxed{\frac{3}{v^3} = \frac{2}{v_{\perp}^3} + \frac{1}{v_{\parallel}^3}}. \quad (40.22)$$

### Характеристическая температура Дебая

Подставляя (40.21) и (40.8) в выражение (40.13) для внутренней энергии, мы получаем

$$U = V \frac{3}{2\pi^2 v^3} \int_0^{\omega_{\max}} \left( \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\hbar\omega/(k_B T)} - 1} \right) \omega^2 d\omega, \quad (40.23)$$

где  $\omega_{\max}$  — максимальная частота нормальных колебаний, которая определяется из соотношения

$$\int_0^{\omega_{\max}} dN = 3N, \quad (40.24)$$

так как полное число нормальных колебаний равно общему числу степеней свободы системы. Используя (40.21), находим:

$$3 \frac{N}{V} = 3n = \frac{3}{2\pi^2 v^3} \int_0^{\omega_{\max}} \omega^2 d\omega = \frac{\omega_{\max}^3}{2\pi^2 v^3}, \quad (40.25)$$

где  $n$  — концентрация атомов (их число в единице объема кристалла). Таким образом, максимальная частота нормальных колебаний (называемая *дебаевской частотой*) равна:

$$\omega_{\max} = v \sqrt[3]{6\pi^2 n}. \quad (40.26)$$

Следует отметить, что наименьшая длина упругой волны в кристалле, которая соответствует максимальной частоте  $\omega_{\max}$ , равна:

$$\lambda_{\min} = \frac{2\pi v}{\omega_{\max}} = a \left( \frac{4\pi}{3} \right)^{1/3} \approx 1,6a, \quad (40.27)$$

где  $a = 1/\sqrt[3]{n}$  — расстояние между соседними атомами в кристаллической решетке. Этот результат согласуется с тем, что волны, длины которых меньше удвоенного межатомного расстояния, не могут существовать в кристалле.

Используя определение (40.26) и учитывая, что для одного моля кристалла концентрация атомов равна  $n = n_a(N_A/V)$ , где  $n_a$  — число атомов в молекуле

вещества кристалла, мы можем записать внутреннюю энергию одного моля в виде

$$U_m = 9n_a N_A \int_0^{\omega_{\max}} \left( \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\hbar\omega/(k_B T)} - 1} \right) \frac{\omega^2 d\omega}{\omega_{\max}^3}. \quad (40.28)$$

Дифференцируя внутреннюю энергию  $U_m$  по температуре, получаем молярную теплоемкость кристалла:

$$c_m = 9n_a R \left( \frac{\hbar\omega_{\max}}{k_B T} \right)^2 \int_0^{\omega_{\max}} \frac{e^{\hbar\omega/(k_B T)}}{(e^{\hbar\omega/(k_B T)} - 1)^2} \frac{\omega^4 d\omega}{\omega_{\max}^5}. \quad (40.29)$$

Введем новый параметр — *характеристическую температуру Дебая*:

$$\theta_D = \frac{\hbar\omega_{\max}}{k_B} \quad (40.30)$$

и выполним в интеграле (40.29) замену переменных  $x = \hbar\omega/(k_B T)$ . Тогда молярную теплоемкость кристалла можно записать в виде

$$c_m = 9n_a R \left( \frac{T}{\theta_D} \right)^3 \int_0^{\theta_D/T} \frac{e^x}{(e^x - 1)^2} x^4 dx. \quad (40.31)$$

При низких температурах  $T \ll \theta_D$  верхний предел интеграла будет очень большим, так что его можно приближенно считать равным бесконечности. Тогда интеграл будет представлять собой число

$$\int_0^{\infty} \frac{e^x}{(e^x - 1)^2} x^4 dx = \frac{4\pi^4}{15}$$

и теплоемкость окажется пропорциональной кубу температуры:

$$c_m \approx \frac{12}{5} \pi^4 n_a R \left( \frac{T}{\theta_D} \right)^3, \quad T \ll \theta_D. \quad (40.32)$$

Эта приближенная зависимость известна как **закон Дебая** и хорошо согласуется с экспериментом при достаточно низких температурах.

При высоких температурах  $T \gg \theta_D$  экспонента в числителе приближенно равна единице, а экспоненту в знаменателе можно разложить в ряд Тейлора:  $e^x \approx 1 + x$ . Тогда для молярной теплоемкости получается значение

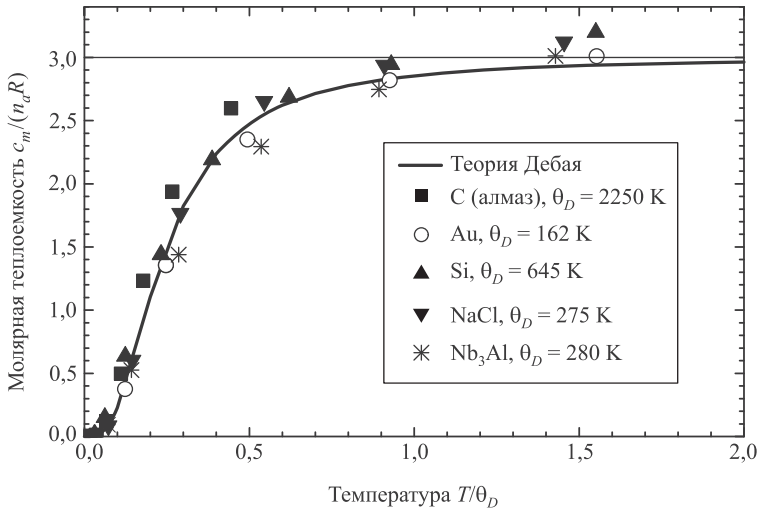
$$c_m \approx 9n_a R \left( \frac{T}{\theta_D} \right)^3 \int_0^{\theta_D/T} x^2 dx = 3n_a R, \quad T \gg \theta_D, \quad (40.33)$$

т. е. закон Дюлонга и Пти.

О согласии теории Дебая с опытом можно судить по графику на рис. 40.3, на котором показаны экспериментальные точки для некоторых веществ.

**Задача 40.61.** Пользуясь данными, приведенными на рис. 40.3, найти максимальную частоту колебаний  $\omega_{\max}$  в кристалле золота по теории Дебая.





**Рис. 40.3.** Сравнение теории теплоемкости Дебая с экспериментальными данными. Показаны вещества с заметно отличающимися значениями дебаевской температуры и разным составом молекул ( $n_a = 2$  для NaCl и  $n_a = 4$  для Nb<sub>3</sub>Al), но все точки лежат достаточно близко к теоретической кривой

**Решение.** Температура Дебая для золота, как указано на графике, равна  $\theta_D = 162$  К. Используя (40.30), находим:

$$\omega_{\max} = \frac{k_B \theta_D}{\hbar} = \frac{1,38 \cdot 10^{-23} \times 162}{1,05 \cdot 10^{-34}} = 2,1 \cdot 10^{13} \text{ Гц.} \blacksquare$$

## 40.5 Фононы

Модель Дебая рассматривает кристалл как сплошную среду, пренебрегая его атомарной структурой. Это означает, что модель пригодна для описания колебаний кристалла в длинноволновом приближении, когда длина волны много больше постоянной решетки. В этом разделе мы изучим колебания кристаллической решетки, рассматривая ее как систему связанных гармонических осцилляторов.

В разделе 24.4 мы рассмотрели колебания двух одинаковых осцилляторов, связанных пружиной. При отсутствии связи осцилляторы колеблются независимо с одинаковой частотой. При наличии связи колебания каждого осциллятора уже не будут гармоническими, но их можно представить как суперпозицию гармонических (так называемых *нормальных*) колебаний с разными частотами. Нормальные моды — это коллективные колебания: в одном случае осцилляторы колеблются синфазно, в другом — в противофазе. Число нормальных частот, на которые «расщепляется» частота колебаний  $N$  независимых осцилляторов при «включении» связи между ними, равно их количеству. Если представить кристалл макроскопических размеров как систему связанных осцилляторов, то при их огромном числе «зазор» между расщепившимися частотами настолько мал, что нормальные частоты можно считать изменяющимися непрерывно в

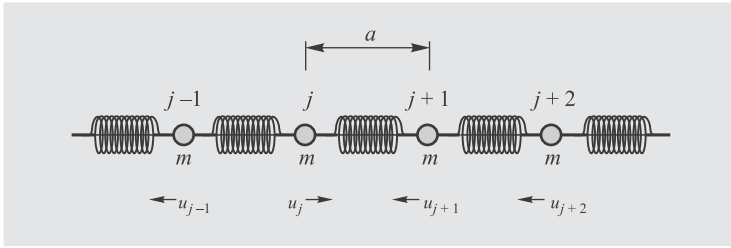


Рис. 40.4. Линейная цепочка одинаковых атомов, связанных пружинами

определенном интервале. Физически нормальные колебания представляют собой монохроматические плоские волны деформаций кристаллической решетки (т. е. звуковые волны). Движение атомов в кристалле может быть описано суперпозицией плоских волн различной частоты.

### Линейная цепочка одинаковых атомов

Чтобы выявить основные закономерности колебаний кристаллической решетки, мы рассмотрим две простейшие системы.

Рассмотрим сначала модель одномерного кристалла, представляющего собой бесконечную цепочку одинаковых материальных точек (атомов) массой  $m$ , связанных между собой пружинами одинаковой жесткостью  $k$  и длиной  $a$  в нерастянутом состоянии (рис. 40.4). Смещение атома с номером  $j$  из положения равновесия обозначим через  $u_j$ . Тогда растяжение пружины между атомами с номерами  $j$  и  $j + 1$  равно  $u_{j+1} - u_j$ , что, согласно закону Гука, приводит к появлению упругой силы  $k(u_{j+1} - u_j)$ , действующей на атом  $j$  со стороны атома  $j + 1$ . Аналогичная сила  $k(u_j - u_{j-1})$  действует в противоположном направлении на атом  $j$  со стороны атома  $j - 1$ . Тогда уравнение движения каждого атома имеет вид

$$m\ddot{u}_j = k(u_{j+1} - u_j) - k(u_j - u_{j-1}) = -k(2u_j - u_{j+1} - u_{j-1}). \quad (40.34)$$

Частота колебаний отдельного (несвязанного) осциллятора равна  $\omega_0 = \sqrt{k/m}$ .

Уравнение плоской волны вдоль цепочки имеет обычный вид

$$u(x, t) = q_K \exp[-i\omega(K)t + iKx]$$

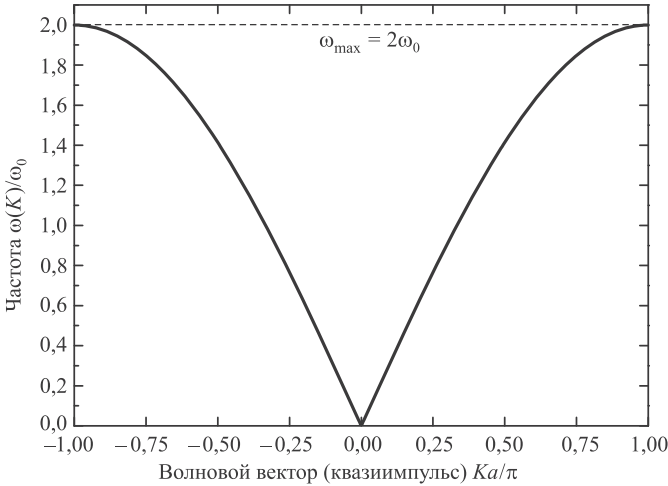
с той разницей, что координата  $x$  пробегает дискретный ряд значений. В кристалле бессмысленно говорить о смещении среды между узлами, физический смысл имеет лишь смещение  $u_j$  узла, находящегося в точке  $x = ja$ . Отсюда следует, что  $\ddot{u}_j = -\omega^2(K)u_j$  и что смещения соседних атомов сдвинуты по фазе относительно друг друга:  $u_{j+1} = \exp(iKa)u_j$  и  $u_{j-1} = \exp(-iKa)u_j$ . Подставляя эти выражения в (40.34), получаем уравнение (и его решение) для нормальных частот колебаний цепочки атомов:

$$\omega^2(K) = 2\omega_0^2 [1 - \cos(Ka)] = 4\omega_0^2 \sin^2(Ka/2);$$

$$\boxed{\omega(K) = 2\omega_0 \left| \sin \left( \frac{Ka}{2} \right) \right|}. \quad (40.35)$$

Смещение каждого узла есть суперпозиция нормальных колебаний:

$$u_j = \sum_K q_K \exp[-i\omega(K)t + iKja]. \quad (40.36)$$



**Рис. 40.5.** Дисперсия акустических колебаний цепочки атомов

Важное следствие: амплитуды нормальных колебаний  $q_K$  (их называют *обобщенными координатами*) не зависят от номера ячейки  $j$ .

Величина  $K$ , называемая *квазиимпульсом*, является аналогом модуля волнового вектора. Из-за периодичности кристалла играют роль лишь его значения в так называемой *первой зоне Бриллюэна*, т.е. в интервале  $-\pi/a \leq K \leq \pi/a$ . Другие значения квазиимпульса ведут к тем же частотам нормальных колебаний. Дисперсия нормальных колебаний, т.е. зависимость их частоты от квазиимпульса, показана на рис. 40.5. Нормальные колебания не есть колебания какого-то из атомов решетки. Это коллективный эффект, синхронизированные колебания всех атомов сразу, это звуковая волна, распространяющаяся с некоторым волновым числом (квазиимпульсом)  $K$  и зависящей от него частотой  $\omega(K)$ . Групповая скорость звуковой волны в этом кристалле равна:

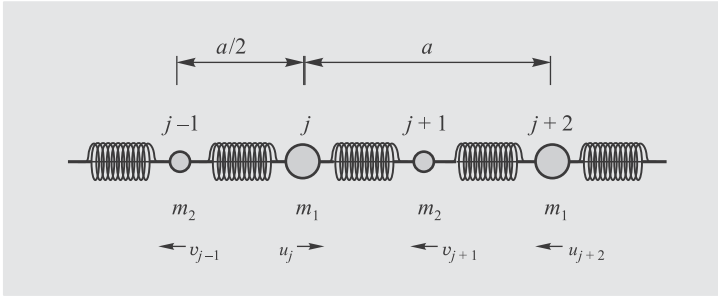
$$v_{\text{гр}} = \frac{d\omega(K)}{dK} = a\omega_0 \cos\left(\frac{Ka}{2}\right). \quad (40.37)$$

На краях первой зоны Бриллюэна квазиимпульс  $K = \pm\pi/a$ . При таком значении волнового числа соседние атомы движутся в противофазе, а атомы, разделенные промежутком  $2a$ , — синфазно, т.е. длина волны  $\lambda = 2\pi/K = 2a$ . Волну с меньшей длиной в такой цепочке возбудить невозможно, так что частота достигает максимального значения  $\omega_{\text{max}} = 2\omega_0$ , а групповая скорость равна нулю (стоячая волна).

Вблизи дна зоны, когда  $|Ka/\pi| \ll 1$ , из уравнения (40.35) следует линейная зависимость частоты от квазиимпульса, характерная для электромагнитных волн (фотонов):  $\omega(K) = \omega_0 Ka$  (см. среднюю часть графика на рис. 40.5). Величина  $a\omega_0$  есть групповая скорость волны в длинноволновом приближении, т.е. скорость распространения звука в нашей цепочке. Такие колебания называют *акустическими*.

### Линейная двухатомная цепочка

Рассмотрим теперь более сложную задачу о нормальных колебаниях линейной цепочки, состоящей из атомов двух сортов, отличающихся массами (рис. 40.6). Смещение атомов с массой  $m_1$  мы по-прежнему обозначим символом  $u_j$ , а смещение атомов с массой  $m_2$  — символом  $v_j$ . Для простоты расстояния между атомами считаем одинаковыми и



**Рис. 40.6.** Линейная цепочка разных атомов; длина пружин в нерастянутом состоянии равна  $a/2$ , так что постоянная этой «решетки» по-прежнему равна  $a$

равными  $a/2$ , так что период этой одномерной «решетки» по-прежнему  $a$ . Жесткость  $k$  всех пружин тоже полагаем одинаковой. В этой модели упругая сила, действующая на атомы одного сорта, зависит от смещения атомов другого сорта, его соседей. Уравнения движения атомов обоих сортов имеют вид

$$\begin{aligned} m_1 \ddot{u}_j &= k(v_{j+1} - u_j) - k(u_j - v_{j-1}) = -k(2u_j - v_{j+1} - v_{j-1}), \\ m_2 \ddot{v}_{j+1} &= k(u_{j+2} - v_{j+1}) - k(v_{j+1} - u_j) = -k(2v_{j+1} - u_{j+2} - u_j). \end{aligned} \quad (40.38)$$

Здесь существуют две частоты колебаний несвязанных осцилляторов:  $\omega_{01} = \sqrt{k/m_1}$  и  $\omega_{02} = \sqrt{k/m_2}$ .

Нормальные колебания (плоские волны) имеют две ветви с разными амплитудами:

$$\begin{aligned} u_j &= \sum_K q_K \exp[-i\omega(K)t + iKja/2], \\ v_{j+1} &= \sum_K Q_K \exp[-i\omega(K)t + iK(j+1)a/2]. \end{aligned} \quad (40.39)$$

Колебания атомов одного сорта сдвинуты по фазе на ту же величину, что и раньше:  $u_{j+2} = \exp(iKa)u_j$ . Подставляя теперь эти соотношения в (40.38), получаем систему уравнений для частот нормальных колебаний

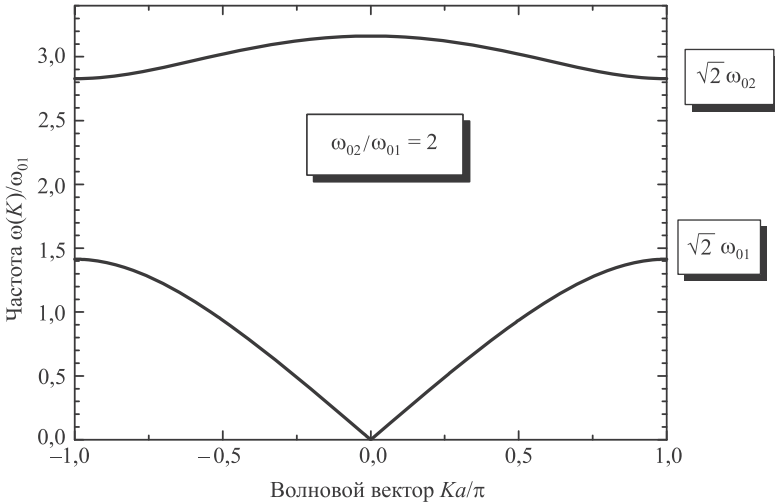
$$\begin{aligned} \omega^2(K)q_K &= 2\omega_{01}^2 \left( q_K - Q_K \cos \frac{Ka}{2} \right), \\ \omega^2(K)Q_K &= 2\omega_{02}^2 \left( Q_K - q_K \cos \frac{Ka}{2} \right). \end{aligned} \quad (40.40)$$

Записываем эти уравнения в привычном виде системы линейных уравнений относительно обобщенных координат  $q_K$  и  $Q_K$ :

$$\begin{aligned} q_K (\omega^2(K) - 2\omega_{01}^2) + 2Q_K \omega_{01}^2 \cos \frac{Ka}{2} &= 0, \\ 2q_K \omega_{02}^2 \cos \frac{Ka}{2} + Q_K (\omega^2(K) - 2\omega_{02}^2) &= 0. \end{aligned} \quad (40.41)$$

Эта система имеет ненулевые решения лишь при равенстве нулю детерминанта, откуда получаем уравнение для частот нормальных колебаний

$$(\omega^2(K) - 2\omega_{01}^2) (\omega^2(K) - 2\omega_{02}^2) - 4\omega_{01}^2 \omega_{02}^2 \cos^2 \frac{Ka}{2} = 0. \quad (40.42)$$



**Рис. 40.7.** Акустическая (нижняя кривая) и оптическая (верхняя кривая) моды колебаний в линейной двухатомной цепочке

Получившееся отсюда биквадратное уравнение

$$\omega^4(K) - 2\omega^2(K)(\omega_{01}^2 + \omega_{02}^2) + 4\omega_{01}^2\omega_{02}^2 \sin^2 \frac{Ka}{2} = 0 \quad (40.43)$$

легко можно решить, что приводит к двум ветвям частот нормальных колебаний  $\omega_+(K)$  и  $\omega_-(K)$ :

$$\omega_{\pm} = \sqrt{\omega_{01}^2 + \omega_{02}^2 \pm \sqrt{\omega_{01}^4 + \omega_{02}^4 + 2\omega_{01}^2\omega_{02}^2 \left(1 - 2\sin^2 \frac{Ka}{2}\right)}}. \quad (40.44)$$

На рисунке 40.7 показаны *дисперсионные характеристики* для двух ветвей нормальных колебаний линейной двухатомной цепочки для частного случая  $\omega_{02} = 2\omega_{01}$  (или  $m_1 = 4m_2$ ). Нижняя ветвь — это уже знакомые нам акустические колебания. Верхняя ветвь называется *оптической модой* (название связано с тем, что впервые такие колебания были обнаружены по излучению в оптическом диапазоне частот). При  $Ka = \pi$  оптическая ветвь соответствует колебаниям только меньших масс  $m_2$ , а акустическая ветвь — колебаниям только больших масс  $m_1$ . Фактор  $\sqrt{2}$  в предельных значениях нормальных частот появляется из-за удвоения жесткости (пружины действуют с двух сторон).

В  $D$ -мерной кристаллической решетке могут распространяться волны, число типов которых равно  $Dr$ , где  $r$  — число атомов в элементарной ячейке кристалла. Из них  $D$  типов относятся к акустическим колебаниям (частоты которых стремятся к нулю при малых квазиимпульсах), остальные же  $D(r - 1)$  колебаний относятся к оптическим модам. Акустическая мода — это колебания атомов решетки как целое. Оптической ветви соответствуют малые длины волн (большие частоты), при которых колебания соседних атомов происходят преимущественно в противоположных направлениях, что связано с изменением дипольного электрического момента системы и оптическим излучением. Выше мы рассмотрели случаи  $D = 1, r = 1$  (один акустический тип волн) и  $D = 1, r = 2$  (одно акустическое и одно оптическое колебания). В кристалле поваренной соли NaCl

( $D = 3, r = 2$ ) могут существовать шесть типов волн — три акустические и три оптические моды. В одноатомных кристаллах ( $r = 1$ ) существуют лишь акустические волны.

### Квантование колебаний кристаллической решетки

Вспомним еще раз колебания двух связанных осцилляторов, рассмотренные в главе 24. Полная энергия системы, изображенной на рис. 24.12, складывается из двух слагаемых, соответствующих кинетической энергии осцилляторов, и трех слагаемых — потенциальной энергии двух основных пружин и третьей, связывающей колеблющиеся грузы:

$$E = \frac{m\dot{x}_1^2}{2} + \frac{m\dot{x}_2^2}{2} + \frac{k(x_1 + L)^2}{2} + \frac{k(x_2 - L)^2}{2} + \frac{\varkappa(x_2 - x_1 - 2L)^2}{2}. \quad (40.45)$$

В нормальных координатах  $\xi_1$  и  $\xi_2$ , определенных уравнениями (24.54), эта энергия равна сумме энергий двух несвязанных осцилляторов:

$$E = \left( \frac{m\xi_1^2}{2} + \frac{k\xi_1^2}{2} \right) + \left( \frac{m\xi_2^2}{2} + \frac{(k + 2\varkappa)\xi_2^2}{2} \right). \quad (40.46)$$

В квантовой механике энергию такой системы также можно записать как сумму энергий двух квантовых осцилляторов с разными частотами  $\omega_1 = \sqrt{k/m}$  и  $\omega_2 = \sqrt{(k + 2\varkappa)/m}$ :

$$E_{n_1, n_2} = \hbar\omega_1 \left( n_1 + \frac{1}{2} \right) + \hbar\omega_2 \left( n_2 + \frac{1}{2} \right). \quad (40.47)$$

Два набора квантовых чисел ( $n_1 = 0, 1, 2, \dots$  и  $n_2 = 0, 1, 2, \dots$ ) полностью определяют состояние системы; мы также можем выписать волновые функции в нормальных координатах.

В кристаллах мы тоже имеем сложные колебания атомов, которые могут быть представлены как суперпозиция монохроматических волн с волновым вектором  $\vec{K}$  и частотами  $\omega_i(\vec{K})$ , где индекс  $i$  нумерует тип волны ( $i = 1, 2, \dots, 3r$ ). Следуя идеям де Бройля, каждой такой волне можно сопоставить частицу с энергией  $\hbar\omega_i(K)$  и импульсом  $\vec{P} = \hbar\vec{K}$ . Эта квазичастица носит название **фонона**. Фонон — элементарная порция звуковой энергии, подобно тому как фотон — элементарная порция световой (электромагнитной) энергии. Наше сопоставление можно схематически изобразить следующим образом:

звуковая (акустическая) волна  $\implies$  квантовый осциллятор  $\implies$  фонон.

При квантово-механическом рассмотрении гармонический осциллятор данной моды может иметь энергию  $E_{i, n_i} = \hbar\omega_i(K)(n_i + \frac{1}{2})$ , где  $n_i$  пробегает целочисленные значения:  $n_i = 0, 1, 2, \dots$ . Вкупе с гипотезой фононов это приводит к естественной интерпретации  $n_i$  как числа фононов данной моды. При  $n_i = 0$  мы имеем нулевые колебания с энергией  $E_{0,0} = \hbar\omega_i(K)/2$  — фонона данной моды в твердом теле нет. При  $n_i = 1$  мы имеем новое состояние с энергией возбуждения  $E_{i,1} - E_{i,0} = \hbar\omega_i(K)$  — это и есть квазичастица фонон. При произвольном

квантовом числе  $n_i$  энергия возбуждения равна  $E_{i,n_i} - E_{i,0} = n_i \hbar \omega_i(K)$ . В таком случае мы говорим, что в твердом теле распространяются  $n_i$  фононов данной моды  $i$ . Полную энергию колебаний находим как сумму энергий всех фононных мод. Иными словами, коллективные колебания кристаллической решетки можно рассматривать как газ свободных или слабо взаимодействующих фононов.

С использованием полученных выше результатов в случае термодинамического (теплого) равновесия можно найти среднее число фононов  $\langle n_i \rangle$  с частотой  $\omega_i(K)$ . Действительно, в (40.7) мы уже нашли среднее значение колебательного числа для квантового осциллятора, и нам осталось заменить частоту  $\omega$  на частоту волны  $\omega_i(K)$ :

$$\langle n_i \rangle = \frac{1}{\exp\left(\frac{\hbar \omega_i(K)}{k_B T}\right) - 1}. \quad (40.48)$$

При низких температурах  $k_B T \ll \hbar \omega_i(K)$ , где важны квантовые эффекты, среднее число оптических фононов экспоненциально убывает: тепловой энергии не хватает на возбуждение колебаний с высокой частотой. Акустических фононов большой энергии также очень мало. Приближенно можно считать, что для частот  $\omega_i > \omega_{\max}$  их вовсе нет, а закон дисперсии мало отличается от линейного, и тогда мы приходим к модели Дебая.

Наоборот, при высоких температурах  $T \gg \hbar \omega_i(K)$  экспоненту в знаменателе можно разложить в ряд Тейлора и получить результат  $n_i \approx \frac{k_B T}{\hbar \omega_i(K)}$ . Следовательно, из полученного соотношения вытекает, что при достаточно высокой температуре в кристалле может одновременно возбуждаться неограниченное количество одинаковых фононов, т.е. принцип Паули на фононы не распространяется. Напомним, что кванты электромагнитного поля — фотоны, находящиеся в состоянии равновесия со стенками полости, — также подчиняются этому распределению. Стало быть, как фонон, так и фотон относятся к одному классу частиц — к **бозонам** (с бозонами и фермионами мы уже встречались в разделе 37.6, когда речь шла о принципе Паули).

Представление о фононах широко используется в физике твердого тела. Фононы — один из примеров так называемых **квазичастиц**. Квазичастицы не менее реальны, чем обычные частицы, но они могут существовать только в среде. Фонон не может вылететь за пределы кристалла, так как сам по себе является коллективным возбуждением кристаллической решетки. В остальном они вполне реальны. Фононы обладают энергией и (квази)импульсом, можно определить их эффективную массу. Фононы могут взаимодействовать друг с другом и с обычными частицами (электрон-фононное взаимодействие в кристалле). Закон дисперсии (имеется в виду связь энергии и импульса) для обычных частиц неизменен: квадратичный для нерелятивистских частиц и линейный для фотонов. Закон дисперсии фононов (квазичастиц) зависит от кристаллической решетки и может иметь весьма замысловатый вид (см. рис. 40.5, 40.7).

Идея квазичастиц впервые была выдвинута Л. Д. Ландау в 1940-х гг. Кроме фононов, есть и другие типы квазичастиц: *магноны* (элементарные возбуждения в системе взаимодействующих спинов), *плазмоны* (коллективные колебания

электронов в плазме), *полярны* (квазичастицы, связанные с облаком поляризации, образующимся при движении электрона в диэлектриках) и др. Тепловые колебания решетки можно рассматривать как фононный газ, при низких температурах — идеальный. При высоких температурах решетка плавится и модель невзаимодействующих фононов неприменима: они перестают быть свободными. Преимущество представления о фононах состоит в том, что в его рамках свойства твердого тела рассматриваются как свойства ансамбля большого числа независимых квазичастиц — фононного газа. Все представления этой модели могут быть использованы для описания поведения кристаллической решетки.

Можно рассматривать также взаимодействие обычных частиц (электронов, фотонов) с фононами. Так, электроны обмениваются фононами, что порождает притяжение. Несмотря на кулоновское отталкивание, может даже образоваться связанное состояние пары электронов. Подобный механизм ведет к явлению *сверхпроводимости*.

Ранее (см. разд. 38.4) мы обсуждали комбинационное рассеяние света кристаллами. Этот процесс можно трактовать как процесс взаимодействия фотона с фононами. Фотон, пролетающий через кристаллическую решетку, может возбудить в ней фонон одной из частот *оптической моды* кристалла. При этом фотон расходует часть своей энергии, вследствие чего его частота уменьшается — возникает красный спутник. Если в кристалле уже был возбужден фонон, то пролетающий фотон может поглотить его, увеличив при этом свою энергию; в таком случае возникает фиолетовый спутник.

Неупругое взаимодействие света с *акустическими фононами* называют *рассеянием Манделъштама—Бриллюэна*. В обычных объемных материалах доминируют процессы взаимодействия света с оптическими фононами. Но применение в электронике наноструктур — нульмерных квантовых точек, одномерных квантовых проволочек и двумерных квантовых пленок, где в процесс вовлекаются фононы с большими квазимпульсами, — привело к повышению роли взаимодействия света с акустическими фононами.

Мы уже убедились, что число фононов в твердом теле не постоянно. Фононов тем больше, чем интенсивнее тепловое движение атомов, т. е. чем выше температура. При высоких температурах число фононов пропорционально температуре, а с приближением к абсолютному нулю их число стремится к нулю.

## 40.6 Статистика Бозе—Эйнштейна

Как мы продемонстрировали в разделе 37.7, *бозоны* и *фермионы* имеют совершенно разные статистические свойства, т. е. по-разному ведут себя в коллективе себе подобных. Кроме непосредственного силового взаимодействия частиц, имеется специфически квантовое, обменное взаимодействие: это не какие-то дополнительные силы или поля — одни частицы влияют на поведение других своим присутствием. Эти эффекты ощущаются частицами, если они находятся друг от друга на расстояниях, меньших или порядка длины волны де Бройля  $\lambda$ . При высоких температурах энергии частиц велики и длина волны де Бройля много меньше расстояния между частицами — это область классической физики. При низких температурах  $\lambda$  возрастает и квантовые эффекты доминируют.



Рассмотрим систему одинаковых бозонов с энергиями  $E_i$  в состоянии  $i$  (где  $i$  обозначает полный набор квантовых чисел). Обозначим через  $n_i$  число частиц в состоянии  $i$  (весь набор  $\{n_i\}$  называется **числами заполнения**). Основной принцип статистической физики (классической и квантовой): вероятность обнаружить систему в таком состоянии равна:

$$W = A \exp \left[ - \sum_i n_i \frac{E_i - \mu}{k_B T} \right], \quad (40.49)$$

где  $A$  — нормировочная постоянная, а  $\mu$  — **химический потенциал**. Сумма  $\sum_i n_i E_i$  есть полная энергия системы, так что по сути уравнение (40.49) обобщает известное распределение Больцмана (см. т. 1, разд. 13.4) на случай  $N$  частиц, энергия которых принимает дискретные значения. Для полного сходства множитель с химическим потенциалом в экспоненте можно было бы включить в нормировку, но он нам важен в явном виде, как мы скоро убедимся<sup>1</sup>.

Для бозонов число  $n_i$  частиц в состоянии  $i$  может быть любым от 0 до  $\infty$  (или от 0 до  $N$  при фиксированном числе частиц). Рассмотрим какое-то конкретное состояние  $k$  системы, энергия которого равна  $E_k$ . Тогда для вероятности того, что в этом состоянии окажется  $n$  частиц, получаем из основного соотношения (40.49) выражение

$$W_k(n) = C_k \exp \left[ -n \frac{E_k - \mu}{k_B T} \right], \quad (40.50)$$

где мы ввели обозначение  $C_k = A \sum_{i \neq k} \exp[n_i(\mu - E_i)/(k_B T)]$ . Сумма  $W_k(n)$  по всем значениям  $n$  (включая и нулевое) равна вероятности того, что в состоянии  $k$  окажется какое-то количество частиц или не окажется ни одной. Очевидно, такая сумма должна быть равна единице:

$$\sum_{n \geq 0} W_k(n) = C \sum_{n \geq 0} \exp \left[ -n \frac{E_k - \mu}{k_B T} \right] = \frac{C_k}{1 - \exp \left( -\frac{E_k - \mu}{k_B T} \right)} = 1. \quad (40.51)$$

Здесь мы в очередной раз использовали формулу для суммы геометрической прогрессии. Теперь легко найти нормировочную постоянную

$$C_k = 1 - \exp \left( -\frac{E_k - \mu}{k_B T} \right) \quad (40.52)$$

и полное выражение для вероятности  $W_k(n)$ . Нас интересует в первую очередь среднее число частиц в состоянии  $k$ , которое по смыслу вероятностей выражается как  $\langle n_k \rangle = \sum_{n \geq 0} n W_k(n)$ . Сумму ряда вида  $\sum_{n \geq 1} n q^n$  мы тоже уже вычисляли.

<sup>1</sup>Химический потенциал всегда появляется в уравнениях статистической физики при фиксированном числе частиц в системе, которое равно  $N = \sum_i n_i$ , где сумма берется по всем возможным состояниям. Математически химический потенциал возникает как множитель Лагранжа при нахождении наиболее вероятного распределения частиц (экстремума статистического веса) при дополнительном условии постоянства числа частиц.

В результате приходим к основному соотношению *квантовой статистики Бозе—Эйнштейна*<sup>2</sup>:

$$\langle n_k \rangle = \frac{1}{\exp\left(\frac{E_k - \mu}{k_B T}\right) - 1}. \quad (40.53)$$

Можно избавиться от индекса  $k$ , записав среднее число частиц в состоянии с энергией  $E$  в виде

$$\langle n_E \rangle = \frac{1}{\exp\left(\frac{E - \mu}{k_B T}\right) - 1}. \quad (40.54)$$

Эта величина называется также *функцией распределения Бозе*.

По физическому смыслу  $n_E \geq 0$ , следовательно,  $\mu < E_0$ , где  $E_0$  — минимальная энергия системы (т. е. энергия основного уровня). Для свободных частиц  $E_0$  равна нулю. Отсюда следует, что для свободных бозонов химический потенциал отрицателен. Его значение зависит от числа частиц в системе, уравнение связи  $\mu$  и  $N$  в случае фиксированного числа частиц имеет вид

$$N = \sum_k \frac{1}{\exp\left(\frac{E_k - \mu}{k_B T}\right) - 1}. \quad (40.55)$$

При высоких температурах достаточно оставить только первый член в разложении экспоненты в ряд, и тогда  $n_k \propto T$ , т. е. число бозонов в каждом состоянии растет с увеличением температуры. Примером может служить обсуждавшийся в предыдущем разделе рост с температурой числа фононов — квантов тепловых колебаний кристаллической решетки.

Если в системе число частиц не фиксировано, а определяется условиями равновесия (например, равновесие излучения с веществом, когда фотоны поглощаются и испускаются), то  $\mu = 0$  (то же самое справедливо, как мы видели, для ансамбля фононов в твердом теле). Применяя (40.53) к системе фотонов ( $\mu = 0$ ,  $E_k = \hbar\omega_k$ ), приходим к формуле (39.15) и ее аналогу для фононов (см. (40.48)).

Заметим, что при  $\exp[(E - \mu)/(k_B T)] \gg 1$  формула (40.54) переходит в классическое распределение Больцмана. В этом случае  $n_E \ll 1$ , что можно интерпретировать как условие малой плотности частиц — квантовый аналог разреженного газа, когда в каждом квантовом состоянии фактически находится не более одной частицы. Значит, несущественно не только прямое взаимодействие частиц, но и их квантовое влияние друг на друга, обменные эффекты.

<sup>2</sup>Ш. Бозе — индийский физик, один из создателей квантовой статистики. Вывел (1924) формулу Планка для распределения энергии, излучаемой абсолютно черным телом, исходя из предположения, что свет можно рассматривать как идеальный газ фотонов и что два состояния этой системы, отличающиеся перестановкой любых двух фотонов, считаются тождественными. Свою статью послал Эйнштейну, который оценил новый подход и немедленно обобщил результаты Бозе на систему любых частиц. Обе статьи были опубликованы в 1925 г. В честь Бозе частицы с целым спином названы бозонами.

## Плотность квантовых состояний

С самого начала оговорим, что *результаты данного раздела в равной степени справедливы как для бозонов, так и для фермионов*. Здесь используется лишь условие квантования импульса (энергии) частицы, заключенной в потенциальную яму конечного объема, да еще закон дисперсии — зависимость энергии частицы  $E(p)$  от ее импульса. Важно также, что все рассуждения ведутся для идеального газа квантовых частиц, не взаимодействующих ни друг с другом, ни с внешним полем.

В теории твердого тела и статистической физике важную роль играет *плотность квантовых состояний*  $\rho(E)$  — число энергетических состояний, приходящихся в единичном объеме на единичный интервал энергии вблизи значения энергии  $E$ . Число состояний мы на самом деле уже подсчитали, рассуждая о числе стоячих волн в разделе 40.4. В формуле (40.18) для числа типов колебаний (числа состояний) заменим волновой вектор на импульс частицы  $\vec{p} = \hbar\vec{k}$ . Получаем тогда:

$$dN = \frac{gV}{(2\pi\hbar)^3} d^3p = 4\pi gV \frac{1}{(2\pi\hbar)^3} p^2 dp. \quad (40.56)$$

Заменив элемент объема  $d^3p$  импульсного пространства на объем  $4\pi p^2 dp$  шарового слоя между сферами радиусами  $p$  и  $p + dp$ , мы предположили изотропию системы. Напомним, что в случае частиц множитель  $g = 2s + 1$  учитывает число спиновых состояний.

Используем теперь закон дисперсии  $E(p)$ . Энергия свободной нерелятивистской частицы  $E = p^2/2m$ , так что  $pdp = m dE$  и  $p = \sqrt{2mE}$ . Перемножив эти соотношения, подставим получившееся выражение для  $p^2 dp$  в формулу (40.56):

$$dN = 2\pi gV \frac{(2m)^{3/2}}{(2\pi\hbar)^3} \sqrt{E} dE. \quad (40.57)$$

Таким образом, для квантовой частицы мы нашли число ее состояний  $dN$  в интервале энергий  $(E, E + dE)$ . Плотность состояний  $\rho(E)$  определится отсюда как

$$\rho(E) = \frac{1}{V} \frac{dN}{dE} = 2\pi g \frac{(2m)^{3/2}}{(2\pi\hbar)^3} \sqrt{E}. \quad (40.58)$$

С ростом энергии плотность числа квантовых состояний растет.

**Задача 40.62.** Найти плотность квантовых состояний для электронов ( $s = 1/2, g = 2$ ) в тонкой двумерной пленке.

**Решение.** Если движение частиц в двумерной пленке ограничено прямоугольником со сторонами  $L_x, L_y$ , то для числа состояний вместо (40.17) получим

$$dN = g dn_x dn_y = 2 \frac{S}{(2\pi)^2} d^2k = 2 \frac{S}{(2\pi\hbar)^2} d^2p. \quad (40.59)$$

Здесь  $S = L_x L_y$  — площадь, доступная частице. Учитывая изотропию системы, заменим элемент площади  $d^2p$  элементарной ячейки в импульсном пространстве на площадь  $2\pi p dp$  между concentрическими окружностями радиусами  $p$  и  $p + dp$ . Тогда (40.59) можно записать в виде

$$dN = 4\pi \frac{S}{(2\pi\hbar)^2} p dp. \quad (40.60)$$

Из квадратичного закона дисперсии для нерелятивистских частиц следует, как и выше,  $pdp = mdE$ , откуда находим:

$$dN = S \frac{m}{\pi \hbar^2} dE. \quad (40.61)$$

Из (40.61) получаем плотность состояний на единицу площади в единичном интервале энергий:

$$\rho_{2D} = \frac{dN}{SdE} = \frac{m}{\pi \hbar^2}. \quad (40.62)$$

Мы нашли, что плотность квантовых состояний в двумерной пленке не зависит от энергии частицы. ■

Двумерный электронный газ, где частицы могут свободно двигаться лишь в двух направлениях, а в третьем их движение ограничено потенциальной ямой, созданной электрическим полем, формируется в так называемых МОП-структурах (от слов «металл — оксид — полупроводник»). Такие гетероструктуры широко применяются при производстве микросхем и полевых транзисторов.

## Конденсация Бозе—Эйнштейна

Рассмотрим идеальный газ свободных бозонов, которые не взаимодействуют ни друг с другом, ни с внешним полем. Состояние частицы задается ее импульсом  $\vec{p}$  и проекцией спина (имеется всего  $g = 2s + 1$  возможностей, для бозонов спиновое квантовое число, или спин  $s$ , принимает целые значения).

Поскольку принцип запрета Паули на бозоны не распространяется, они могут накапливаться в одном состоянии. При нулевой температуре все частицы системы должны занять нижний уровень с энергией  $E = 0$ . Возникает вопрос: что будет при  $T \neq 0$ ?

Рассмотрим интервал энергий  $(E, E + dE)$ . Число состояний  $dN$  в этом интервале дается выражением (40.57), а среднее число бозонов  $\langle n_E \rangle$  в каждом состоянии дается формулой (40.53), где надо полагать  $E_k = E$ . Умножая  $\langle n_E \rangle$  на  $dN$ , получаем **распределение Бозе—Эйнштейна** для числа частиц  $dN$  в указанном интервале энергий:

$$dN = Vg \frac{m^{3/2}}{\sqrt{2\pi^2 \hbar^3}} \frac{\sqrt{E} dE}{\exp\left(\frac{E - \mu}{k_B T}\right) - 1}. \quad (40.63)$$

Величина  $f(E, T)_B = dN/(VdE)$  называется **функцией распределения бозонов по энергиям**:

$$f_B(E, T) = g \frac{(2m)^{3/2}}{4\pi^2 \hbar^3} \frac{\sqrt{E}}{\exp\left(\frac{E - \mu}{k_B T}\right) - 1}. \quad (40.64)$$

Чтобы найти полное число частиц в системе, проинтегрируем (40.63) по всем значениям энергии:

$$N = V \frac{gm^{3/2}}{\sqrt{2\pi^2 \hbar^3}} \int_0^\infty \frac{\sqrt{E} dE}{\exp\left(\frac{E - \mu}{k_B T}\right) - 1}. \quad (40.65)$$

Вводя новую переменную интегрирования  $z = E/(k_B T)$ , переписываем (40.65) в виде

$$\frac{N}{V} = \frac{g(mk_B T)^{3/2}}{\sqrt{2\pi^2 \hbar^3}} \int_0^\infty \frac{\sqrt{z} dz}{\exp[z - \mu/(k_B T)] - 1}. \quad (40.66)$$

Это — уравнение для химического потенциала  $\mu$ . Если при заданной концентрации частиц  $n = N/V$  понижать температуру газа, то химический потенциал будет увеличиваться (т. е. уменьшаться по модулю, так как  $\mu \leq 0$ ), как следует из (40.66). Предельного значения  $\mu = 0$  он достигнет при температуре  $T_0$ , определяемой уравнением

$$n = \frac{g(mk_B T_0)^{3/2}}{\sqrt{2\pi^2 \hbar^3}} \int_0^\infty \frac{\sqrt{z} dz}{e^z - 1}, \quad (40.67)$$

откуда следует выражение

$$T_0 = 3,3125 \frac{\hbar^2}{mk_B} \left(\frac{n}{g}\right)^{2/3}. \quad (40.68)$$

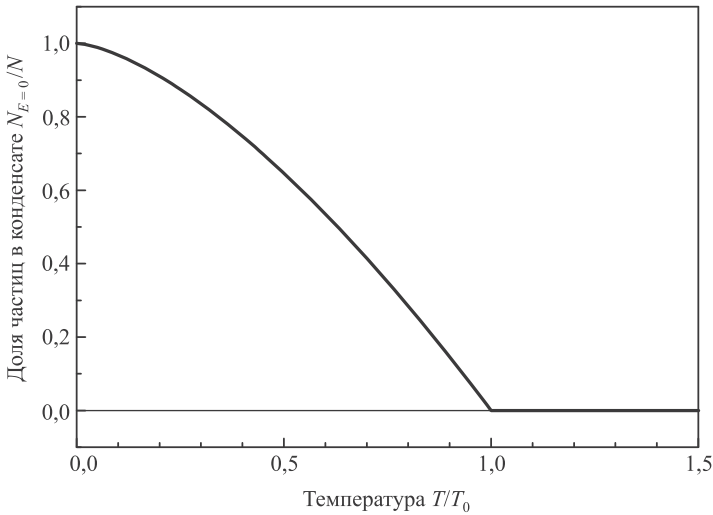
Разберемся в физическом смысле входящих в формулу (40.68) комбинаций параметров. Если обозначить через  $l$  среднее расстояние между частицами, то в объеме радиусом  $l/2$  будет содержаться одна частица, т. е. концентрация частиц  $n \sim 1/l^3$ . Поэтому формулу (40.68) можно переписать в виде  $k_B T_0 \sim \hbar^2/(ml^2)$ . Поскольку тепловая энергия движения частиц  $mv^2 \sim k_B T_0$ , мы получаем отсюда  $l^2 \sim \hbar^2/(m^2 v^2)$ , или  $l \sim \hbar/(mv) = \lambda$ , где  $\lambda$  — длина волны де Бройля (мы опустили в этом рассуждении несущественные числовые множители). Таким образом, снова приходим к выводу, что квантовые эффекты становятся заметными, когда длина волны де Бройля имеет порядок расстояния между частицами. Формула (40.68) — это общее выражение для «квантовой» температуры: оно еще нам встретится, а меняться может лишь числовой множитель.

Итак, при  $T > T_0$  существует физически приемлемое ( $\mu < 0$ ) решение уравнения (40.66). При  $T < T_0$  химический потенциал остается равным предельному значению 0, дальше ему меняться некуда, и правая часть уравнения становится меньше  $N/V$ . Этот результат удивителен, ибо число частиц в системе фиксировано и концентрация частиц, казалось бы, должна оставаться неизменной. Стало быть, какая-то доля этих частиц куда-то девается, выпадает из системы, перестает участвовать в тепловом движении. Значит, правая часть (40.66) будет теперь описывать те частицы, которые участвуют в тепловом движении, т. е. энергия которых больше нуля:

$$N_{E>0} = V \frac{g(mk_B T)^{3/2}}{\sqrt{2\pi^2 \hbar^3}} \int_0^\infty \frac{\sqrt{z} dz}{e^z - 1}, \quad T \leq T_0. \quad (40.69)$$

Если умножить и разделить правую часть (40.69) на  $T_0^{3/2}$ , то можно выделить фактор  $(T/T_0)^{3/2}$ , а оставшийся сомножитель, в соответствии с (40.67), будет равен полному числу  $N$  частиц в системе:

$$N_{E>0} = N \left(\frac{T}{T_0}\right)^{3/2}, \quad T \leq T_0. \quad (40.70)$$



**Рис. 40.8.** Доля частиц в бозе—эйнштейновском конденсате в зависимости от безразмерной температуры  $T/T_0$

Остальные частицы имеют нулевую энергию; их число  $N_{E=0}$  определится как разность  $N - N_{E>0}$ :

$$N_{E=0} = N \left[ 1 - \left( \frac{T}{T_0} \right)^{3/2} \right], \quad T \leq T_0. \quad (40.71)$$

Итак, при температуре ниже критической  $T \leq T_0$  происходит так называемая **бозе—эйнштейновская конденсация** — накопление частиц в состоянии с  $\vec{p} = 0$ . Эффект этот макроскопический, так как число частиц конденсата огромно, порядка полного числа частиц  $N$ . При понижении температуры до критического значения конденсат появляется внезапно — система испытывает резкое изменение свойств (как говорят, происходит фазовый переход).

При абсолютном нуле температуры все частицы системы находятся в конденсате. При повышении температуры некоторые из них покидают конденсат и начинают участвовать в тепловом движении. При  $T > T_0$  бозе—эйнштейновский конденсат исчезает: в нем не остается частиц. Зависимость доли частиц в конденсате от температуры показана на рис. 40.8.

Важно отметить, что конденсация происходит не в обычном пространстве, как при выпадении, например, росы, а в пространстве импульсов. Если говорить в терминах нашего обычного пространства, то и конденсат, и нормальный компонент бозе-газа «размешаны», и система остается пространственно однородной. Но есть и сходство с обычным насыщенным паром: например, при  $T \leq T_0$  давление не зависит от объема системы, что для идеального газа не реализуется.

Явление, теоретически предсказанное еще в середине 1920-х гг., впервые экспериментально наблюдалось в 1995 г. в США. «Газ» из 2000 атомов рубидия, охлажденный почти до абсолютного нуля (170 нК), «сконденсировался» в «суператом», который вел себя уже не как газ частиц, а как единое целое. Чуть

позже удалось получить бозе—эйнштейновский конденсат атомов натрия. Таким образом, было создано новое состояние вещества, никогда ранее не существовавшее в естественных условиях во Вселенной, как было сказано в сообщении о полученных результатах. В 2001 г. авторам этих экспериментальных работ — Э. Корнеллу, В. Кеттерле и К. Виману — была присуждена Нобелевская премия «за достижение конденсации Бозе—Эйнштейна в разреженных газах щелочных металлов и за первые фундаментальные исследования свойств конденсатов». Вскоре были получены бозе—эйнштейновские конденсаты многих изотопов, молекул и квазичастиц.

## Контрольные вопросы

1. Объясните закон Дюлонга и Пти.
2. Перечислите трудности классической теории теплоемкости.
3. В чем основные различия квантовых теорий теплоемкости Эйнштейна и Дебая?
4. Поясните физический смысл характеристической температуры Дебая.
5. Что такое фонон?
6. Опишите различия дисперсионных кривых акустических и оптических фононов.
7. Что такое числа заполнения?
8. Дайте определение плотности квантовых состояний.
9. Что такое функция распределения частиц (бозонов) по энергиям?
10. Что такое конденсация Бозе—Эйнштейна?

## Глава 41

# Элементы физики твердого тела. Статистика Ферми—Дирака

Цель настоящей главы — описать поведение носителей заряда в твердых телах с позиций квантовой механики. Ранее мы обсуждали квантовые состояния электронов в изолированных атомах. Здесь мы изучим, как меняется их поведение, если атомы оказываются в кристаллической решетке. Если рассматривать атом как некую потенциальную яму, захватывающую электроны, то объединение атомов в кристалл приводит к появлению в нем периодического потенциального поля. Оказывается, что движение электронов в таком поле отличается важными особенностями. Электроны, обладающие энергией в некоторых диапазонах, движутся в кристалле свободно, а иметь другие энергии им просто запрещено. Квантовая механика позволила создать стройную теорию электронных свойств твердых тел, объясняющую такие явления и свойства, как электропроводность металлов, различия между металлами, диэлектриками и полупроводниками, их оптические свойства.

### 41.1 Энергетические зоны в твердых телах

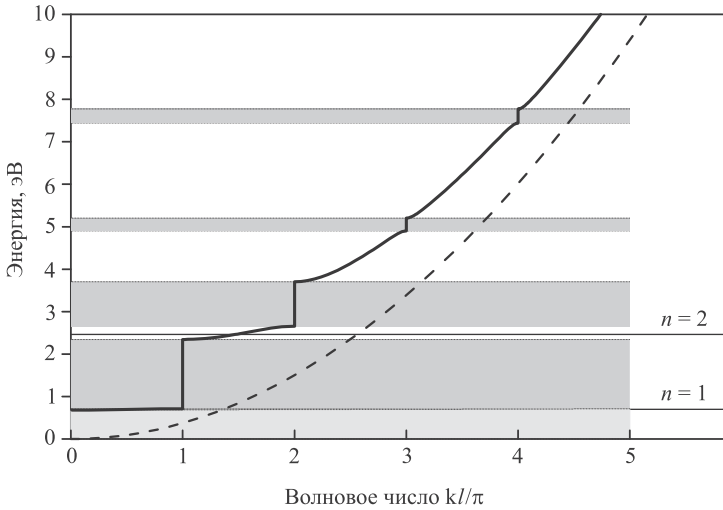
#### Применение оптической аналогии

Вспомним нашу аналогию с оптикой при изучении прохождения света через многослойную структуру. Такая структура, представляющая собой периодический ряд потенциальных барьеров, может служить для моделирования кристаллической решетки. Результаты для частицы мы представим в чуть ином виде. Все формулы (36.62), (36.63), (36.65) остаются в силе, только для  $D_r$  надо взять выражение (36.56). Уравнение же (36.64) можно переписать в виде

$$S = \cos(dk_1) \cos(bk_2) - \frac{k_1^2 + k_2^2}{2k_1k_2} \sin(dk_1) \sin(bk_2). \quad (41.1)$$

Учитывая, что при большом числе ( $N \rightarrow \infty$ ) потенциальных барьеров широкие минимумы коэффициента прохождения  $D^N$  на рис. 36.11 (в пределе  $D^{(\infty)} = 0$ ) получаются из гиперболической ветви решения  $|S| \geq 1$ . Мы возьмем лишь



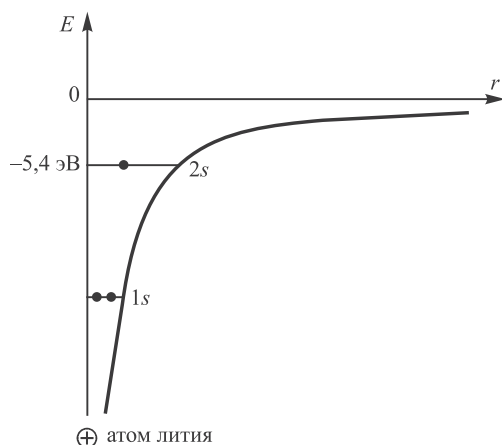


**Рис. 41.1.** Формирование зон при большом числе барьеров. Сплошная линия показывает зависимость энергии от квазиволнового числа (вектора)  $k$ , пунктирная соответствует свободной частице с энергией  $E = \hbar^2 k^2 / (2m)$ . При увеличении энергии сплошная и пунктирные линии сближаются — барьеры оказывают все меньшее влияние, а запрещенные зоны, показанные темным цветом, становятся все уже

тригонометрическую ветвь и решим численно уравнение  $\cos(kl) = S$  относительно энергии частицы  $E = \hbar^2 k_1^2 / (2m)$ , найдя ее зависимость от квазиволнового числа  $k$ . Решение для большого числа барьеров ( $N \rightarrow \infty$ ) представлено на рис. 41.1, где для конкретности мы считаем межбарьерное расстояние равным ширине прямоугольных барьеров:  $d = b = l/2 = 5 \text{ \AA}$ , а высоту барьера  $U = 3 \text{ эВ}$ . При таких значениях изолированная квантовая яма между двумя барьерами имеет два уровня энергии, показанные на рисунке линиями с  $n = 1$  и  $n = 2$ . Нижний уровень практически не изменился из-за того, что мы взяли большое число таких ям: туннелирование между состояниями с  $n = 1$  в соседних ямах маловероятно. Верхние уровни с  $n = 2$  расщепились из-за туннелирования и образовали полосу (зону) разрешенных значений энергии (показана светлым фоном). Выше потенциальных барьеров разрешенные (светлые) зоны соответствуют таким значениям энергии частицы, при которых она не отражается от барьеров. Темные же зоны соответствуют значениям энергии, при которых коэффициент отражения равен единице: частица с такими энергиями не может проникнуть и двигаться даже над барьерами.

### Формирование энергетических зон в кристалле

Здесь мы рассмотрим образование разрешенных и запрещенных зон с несколько иной точки зрения. На рисунке 41.2 показан ход электростатического потенциала в окрестности изолированного атома Li (он помещен в начало координат и обозначен знаком  $\oplus$ ). Горизонтальная линия, помеченная «0», соответствует



**Рис. 41.2.** Низшие энергетические уровни атома лития и характер зависимости электростатического потенциала атома от расстояний  $r$  от ядра

потенциальной энергии, равной нулю, а две более низкие линии — энергетическим уровням  $1s$ - и  $2s$ -состояний. Два электрона локализованы в  $1s$ -состоянии и один — в  $2s$ -состоянии. Для электрона, находящегося в  $2s$ -состоянии, энергия связи равна  $5,4$  эВ; чтобы удалить  $1s$ -электрон, необходимо затратить энергию  $75$  эВ, а чтобы удалить и второй  $1s$ -электрон, нужно еще примерно  $120$  эВ (рисунок дан не в масштабе).

Если атомы лития образуют кристалл, то результирующее электростатическое поле в любой точке внутри кристалла представляет собой сумму полей отдельных атомов. Потенциал между атомами никогда не возрастает до нулевого значения. В действительности потенциал между атомами внутри кристалла существенно ниже потенциала в точке, находящейся на таком же расстоянии от изолированного атома. Это понижение потенциала в кристалле из-за наложения потенциала от соседнего атома настолько значительно, что  $2s$ -электроны, связанные в изолированном атоме, в кристалле уже не встречают потенциального барьера, способного удерживать их в окрестности определенного атома Li. Понижение потенциального барьера освобождает  $2s$ -электроны, волновые функции которых распространяются на весь кристалл. Таким образом, в кристалле лития  $2s$ -электроны ведут себя как свободные электроны (рис. 41.3). Они «принадлежат» уже не отдельным атомам, а кристаллу в целом. Волновые функции  $2s$ -электронов не локализованы, а размазаны по всему кристаллу. Энергетическое  $2s$ -состояние изолированного атома лития превращается в энергетическое «состояние» кристалла. Если в кристалле содержится  $N$  атомов, то все они будут находиться в этом «состоянии». Аналогичная ситуация имеет место для всех кристаллических металлов.

Периодическая решетка кристалла в определенном смысле напоминает систему связанных осцилляторов. В разделе 24.4 мы видели, что частоты нормальных колебаний двух одинаковых слабосвязанных осцилляторов расщепляются на две разные, но близкие частоты. Аналогично частота  $N$  одинаковых связанных осцилляторов расщепится на  $N$  близких частот. Похожим образом  $2s$ -уровни атома лития превратятся в систему близко расположенных уровней кристалла

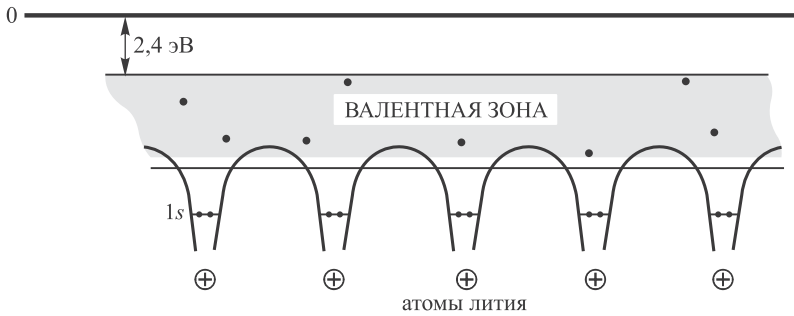


Рис. 41.3. Энергетические уровни и валентная зона кристалла лития

лития. Поскольку число  $N$  атомов в кристалле огромно, разница в энергиях между расщепившимися уровнями энергии настолько мала, что они, по существу, распределены непрерывно, образуя *валентную зону* (см. рис. 41.3; название связано с тем, что эта зона образовалась из-за расщепления уровней, на которых находятся валентные электроны).

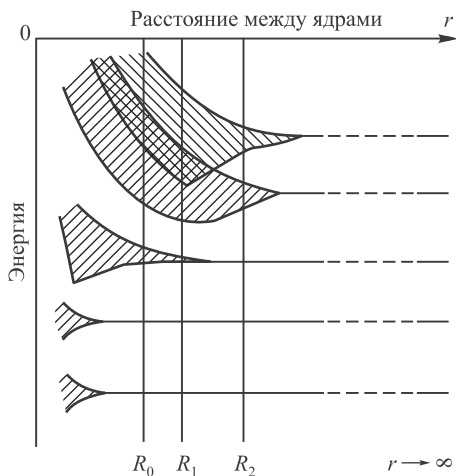
Исходя из представления, что валентные электроны уже не принадлежат отдельным атомам, мы можем сказать, что неопределенность их положения в кристалле характеризуется половиной постоянной решетки (расстояния  $a$  между ее соседними узлами):  $\Delta r_i \sim a/2$ . Тогда по соотношению неопределенности Гейзенберга находим неопределенность компоненты импульса электрона:  $\Delta p_i \sim 2\hbar/a$ . Такой же порядок величины имеет и сама компонента импульса, так что неопределенность  $\Delta E$  энергии электрона равна по порядку величины  $\Delta E \sim \sum_i p_i^2/2m \sim 6\hbar^2/(ma^2)$ . Подставляя значения фундаментальных постоянных, находим:

$$\Delta E \sim \frac{45,7 \text{ эВ}}{(a/1 \text{ \AA})^2}. \quad (41.2)$$

Это соотношение дает оценку ширины валентной зоны. У большинства кристаллов  $a = 2\text{--}4 \text{ \AA}$ , что приводит к значениям  $\Delta E \sim 3\text{--}10 \text{ эВ}$ . Для лития, например,  $a = 3,04 \text{ \AA}$ , так что (41.2) дает  $\Delta E \sim 4,9 \text{ эВ}$ . Мы получили не такую плохую оценку, если учесть, что потолок зоны лежит примерно на 3 эВ выше  $2s$ -уровней (см. рис. 41.2 и 41.3), а из-за упомянутого выше понижения суммарного потенциала дно этой зоны лежит ниже  $2s$ -уровней изолированных атомов лития. Поскольку в одном моле кристалла содержится  $N_A = 6 \cdot 10^{23}$  атомов, то расстояние между соседними уровнями будет равно  $\Delta E/N_A \sim 10^{-23} \text{ эВ}$ , т.е. практически энергия электрона в валентной зоне действительно может принимать непрерывный ряд значений.

Вообще говоря, все энергетические состояния атома превращаются в кристалле в зоны. Пусть первоначально имеется  $N$  изолированных атомов какого-либо вещества. Пока атомы изолированы друг от друга, каждый из них имеет одинаковую систему энергетических уровней. Заполнение уровней электронами осуществляется в каждом атоме независимо от заполнения аналогичных уровней в других атомах. Когда атомы сближаются друг с другом, образуя твердое тело, их

**Рис. 41.4.** Схематическая картина образования зон при сближении атомов. При  $r \rightarrow \infty$  имеется набор энергетических уровней, одинаковый для каждого изолированного атома данного типа. При сближении уровни расширяются, образуя зоны. Нижние уровни возмущаются слабее верхних. Возможно также перекрытие зон (показано двойной штриховкой). Для случая, показанного на рисунке, при равновесном расстоянии  $R_0$  имеется два нижних практически не возмущенных уровня (точнее, они образуют очень узкие зоны), далее следуют изолированная зона и две верхние зоны, слившиеся в одну



взаимодействие оказывает возмущающее действие на первоначальные атомные энергетические уровни. Из каждого уровня изолированного атома (одинакового для всех них) возникают  $N$  очень близких, но не совпадающих уровней. Эти  $N$  близко расположенных друг к другу энергетических уровней образуют зону. Величина расщепления для различных уровней атома неодинакова. Сильнее возмущаются уровни, заполненные внешними электронами, — зоны, образующиеся из них, оказываются широкими. Уровни, заполненные внутренними электронами, возмущаются слабо и образуют очень узкие зоны, мало отличимые от уровней (два нижних уровня на рис. 41.4). При уменьшении расстояния  $r$  между атомами заметно расщепляются лишь уровни, занимаемые валентными электронами (как, например, третий снизу изолированный уровень на рис. 41.4). Более высокие (свободные) уровни, не занятые электронами в невозбужденном состоянии атома, также подвергаются существенному расщеплению и превращению в энергетические зоны, которые даже могут перекрываться.

Некоторые свойства энергетических зон:

- Энергия связи твердого тела обуславливается сдвигом энергетических уровней электронов подобно тому, что происходит при образовании химической связи. Следовательно, при образовании твердого тела энергетические уровни в среднем смещаются вниз.
- Больше всего подвержены возмущающему действию соседних атомов валентные электроны, наиболее удаленные от ядра. Это естественно, так как их связь с «материнским» ионом слабее и они расположены ближе остальных электронов к соседним ионам.
- Равновесное состояние между атомами решетки ( $R_0$  на рис. 41.4) должно соответствовать минимуму энергии, поскольку при дальнейшем сближении атомов энергетические уровни начинают смещаться вверх.
- Состояние исходной системы при сближении атомов должно деформироваться непрерывным образом (см. рис. 41.4). Поэтому число состояний системы  $N$  изолированных атомов не должно изменяться при их объединении

в кристаллическую решетку. Если, например, в кристалл объединяются щелочные металлы, у атомов которых  $s$ -уровни заполнены наполовину, то наполовину будет заполнена и образовавшаяся из них валентная зона.

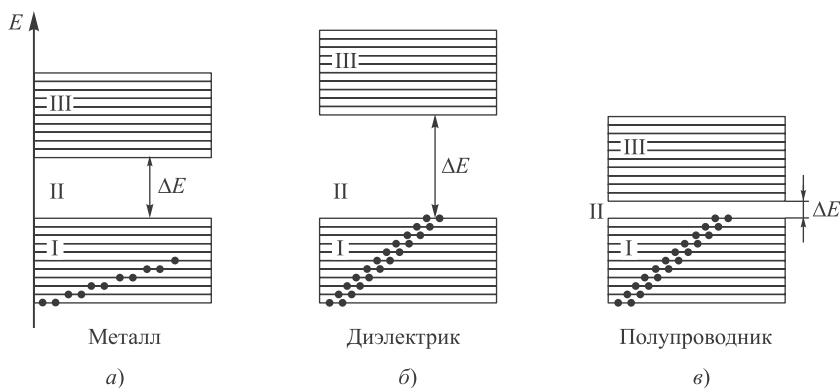
Дискретные атомные уровни проявляются в кристалле либо в виде изолированных зон (когда дно верхней зоны лежит выше верхней границы нижней зоны — как для расстояния  $R_2$  на рис. 41.4), либо в виде перекрывающихся зон, для которых уже нельзя установить, к какой подоболочке первоначально принадлежали уровни (верхние зоны для расстояния  $R_1$  на рис. 41.4). Число уровней в такой слившейся зоне равно сумме количеств уровней, на которые расщепляются оба первоначальных уровня атома. Высшие энергетические зоны всегда перекрываются, в то время как низшие состояния обычно остаются в виде индивидуальных зон кристалла. Перекрытие зон определяется спецификой кристаллической решетки.

В теории твердого тела зонная структура энергетических уровней получается непосредственно из решения уравнения Шрёдингера для электрона, движущегося в периодическом силовом поле кристаллической решетки. Промежутки между энергетическими зонами — это те значения энергии, которые, согласно уравнению Шрёдингера, невозможны в кристалле. Поэтому такие промежутки называют *запрещенными зонами*.

Заполнение энергетических зон электронами происходит по тем же законам, что и заполнение энергетических уровней в атоме. Например, в  $s$ -состоянии могут находиться два электрона, поэтому в кристалле, содержащем  $N$  атомов,  $s$ -зона может содержать  $2N$  электронов, как уже говорилось. Кроме того,  $p$ -зона (если она не перекрывается с  $s$ - и  $d$ -зонами) может содержать  $6N$  электронов и т. д. При целиком заполненной зоне ни один из электронов не может приобрести дополнительную энергию, если только эта величина энергии не окажется достаточной для того, чтобы перевести электрон в незанятое состояние в более высокой зоне. В зависимости от положения различных зон энергия, необходимая для переброски электрона из одной зоны в другую, может составлять 5–10 эВ. Если же высшая энергетическая зона заполнена лишь частично (например,  $2s$ -зона в литии содержит лишь  $N$  электронов и, таким образом, заполнена наполовину), то в этой зоне для находящихся в ней электронов доступно большое число энергетических уровней, расстояние между которыми, как мы видели, чрезвычайно мало. Поэтому электрону частично заполненной зоны можно сообщить любую добавочную энергию, если только его полная энергия не превышает максимальной энергии, допустимой для данной зоны. Такая ситуация характерна и для других щелочных атомов и, как мы увидим далее, должна влиять на их электропроводность.

## 41.2 Металлы, диэлектрики и полупроводники

Энергетические зоны могут быть полностью заполненными электронами, полностью свободными и частично заполненными. В зависимости от конкретной ситуации твердое тело обладает различными свойствами. Существование энергетических зон позволяет объяснить с единой точки зрения разницу свойств *металлов, диэлектриков и полупроводников*.



**Рис. 41.5.** Валентная (I), запрещенная (II) и свободная (III) зоны для металла (а), диэлектрика (б) и полупроводника (в). В первом случае валентная зона заполнена наполовину, в двух других — полностью. Для полупроводника энергетическая щель  $\Delta E$  намного меньше, чем для диэлектрика

При абсолютном нуле температуры валентные электроны с противоположными спинами заполняют попарно (согласно принципу Паули) нижние уровни валентной зоны. В зависимости от степени заполнения валентной зоны электронами и ширины запрещенной зоны возможны три случая, изображенные на рис. 41.5.

Если валентная зона заполнена только частично (рис. 41.5, а), то электроны в ней можно заставить двигаться в определенном направлении, приложив внешнее электрическое поле. В таком случае достаточно сообщить электронам, находящимся на верхних заполненных уровнях валентной зоны, совсем небольшую энергию ( $\sim 10^{-23}$  эВ) для того, чтобы перевести их на более высокие уровни.

Энергия теплового движения ( $k_B T$ ) при комнатных температурах составляет  $1/40$  эВ. Следовательно, значительная часть электронов при комнатной температуре переводится на более высокие энергетические уровни. Дополнительная энергия, вызванная действием на электрон электрического поля, также оказывается достаточной для перевода электрона на более высокие уровни. Поэтому электроны могут ускоряться электрическим полем и приобретать дополнительную скорость в направлении, противоположном направлению поля. Таким образом, кристалл с подобной схемой энергетических уровней будет представлять собой **проводник**.

Валентная зона заполняется частично для таких металлов, как литий или натрий, где на последнем занятом уровне в атоме находится только один электрон (или, более общо, нечетное число электронов, как в атомах алюминия, галлия и некоторых других металлов). Здесь электроны проводимости заполняют попарно только половину уровней валентной зоны.

Когда валентная и свободная зоны перекрываются и образуют единую **зону перекрытия** (ее иногда называют **гибридной зоной**), кристалл может быть металлом и при четном числе валентных электронов. Так, для элементов II группы (бериллий, магний) два электрона находятся в *s*-состоянии, но перекрытие *s*- и *p*-зон обеспечивает наличие свободных уровней для электронов проводимости.

В *диэлектрике* и *полупроводнике* (рис. 41.5, б, в) уровни валентной зоны заполнены полностью. Для того чтобы увеличить энергию электрона, ему необходимо сообщить количество энергии, не меньшее чем ширина  $\Delta E$  запрещенной зоны, отделяющей валентную зону от свободной *зоны проводимости*. Электрическое поле не может заставить двигаться электроны (при условии, что напряженность электрического поля не превышает величины, вызывающей электрический пробой кристалла). Ширина запрещенной зоны определяет электрические свойства кристалла.

Если ширина  $\Delta E$  запрещенной зоны относительно велика (порядка нескольких электронвольт), тепловое движение ( $E \sim k_B T$ ) не сможет забросить в свободную зону значительное число электронов. В этом случае (см. рис. 41.5, б) кристалл оказывается *диэлектриком* (например, в алмазе  $\Delta E \sim 5$  эВ). Если ширина  $\Delta E$  не слишком велика (порядка 0,1–1 эВ), энергия теплового движения оказывается достаточной для того, чтобы перевести часть (хотя бы небольшую) «верхних» электронов в свободную зону (см. рис. 41.5, в). Тогда эти электроны будут вести себя аналогично электронам валентной зоны в металле. Одновременно под действием внешнего электрического поля становится возможным переход «нижних» электронов валентной зоны на ее освободившиеся верхние уровни. Такой кристалл называется *электронным полупроводником*. Например, у самых распространенных полупроводников — кремния и германия — ширина запрещенных зон равна 1,1 и 0,75 эВ соответственно.

### 41.3 Теория свободных электронов в металле. Энергия Ферми

Электроны, движущиеся в металле, притягиваются ионами и не могут вылететь наружу, если их кинетическая энергия не слишком велика. В реальном кристалле максимумы и минимумы потенциальной энергии разделены малыми расстояниями порядка  $10^{-10}$  м. Электрон, прошедший сравнительно малый путь ( $\sim 1$  см) в металле, пересечет области чередования максимумов и минимумов огромное число раз ( $\sim 10^8$ )! Такую быстро колеблющуюся потенциальную энергию можно в первом приближении заменить некоторой средней потенциальной энергией (отрицательной постоянной по всему кристаллу). Другими словами, распределение ионов можно заменить сплошным «размазанным» положительным зарядом. Так как средняя потенциальная энергия постоянна, электрон может перемещаться по кристаллу, не совершая никакой работы (если, разумеется, не выходит за пределы кристалла).

Тепловое движение в металле сводится к колебаниям ионов кристаллической решетки и хаотическому движению свободных электронов. Наличие свободных электронов в металлах подтверждается явлениями проводимости электрического тока и электронной эмиссии. В третьей части первого тома (см. разд. 18.11) мы познакомились с основами классической теории металлов, которая встретила с весьма существенными трудностями.

Во-первых, в классической теории металлов удельное сопротивление  $\rho$  пропорционально среднеквадратичной скорости хаотического теплового движения

электронов (см. формулу (18.64) в т. 1). Кинетическая энергия теплового движения электрона пропорциональна температуре, т.е. его среднеквадратичная скорость пропорциональна квадратному корню из температуры. Следовательно, таким же законом должна, согласно классической теории, описываться зависимость сопротивления от температуры:  $\rho \propto \sqrt{T}$ . Опытные же данные говорят о том, что электрическое сопротивление металлов растет быстрее — пропорционально первой степени температуры.

Во-вторых, согласно классической электронной теории молярная теплоемкость металлов должна быть больше, чем у диэлектриков. В соответствии с законом Дюлонга и Пти (разд. 40.1) теплоемкость одного моля химически простых веществ в кристаллическом состоянии одинакова и равна  $C_m = 3R$ . В металлах электронный газ должен, казалось бы, обладать молярной теплоемкостью, по крайней мере, равной  $C_{m,эл} = 3R/2$ , что должно было бы существенно увеличить теплоемкость кристалла. В действительности же теплоемкость металлов не отличается заметно от теплоемкости диэлектрических кристаллов.

В-третьих, классический результат для теплоемкости  $C_{m,эл} = 3R/2$  противоречит третьему началу термодинамики (см. т. 1, разд. 12.8), согласно которому при абсолютных температурах, стремящихся к нулю, энтропия стремится к нулю, энергия — к постоянному значению и теплоемкость также должна стремиться к нулю.

Объяснение таких несоответствий смогла дать лишь квантовая теория металлов, которая также прошла большой и нелегкий путь развития.

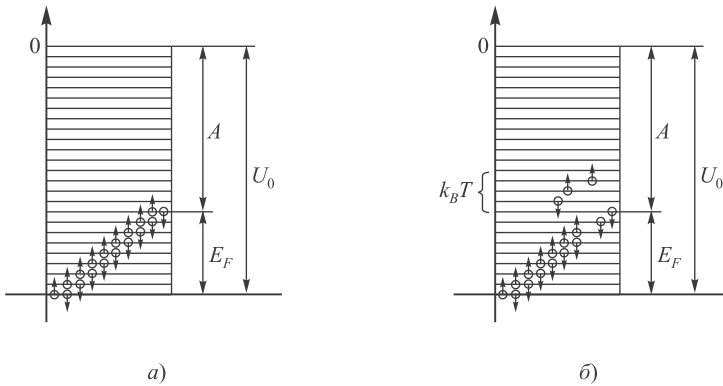
Согласно модели свободных электронов (А. Зоммерфельд<sup>1</sup>, 1927–1928), валентные электроны атомов металла могут почти свободно перемещаться в пределах кристалла. Эти валентные электроны обуславливают электропроводность металла, и по этой причине их называют электронами проводимости.

Мы примем следующую упрощенную модель металла. Будем считать, что имеется лишь одна зона, в которой валентные электроны удерживаются в потенциальной яме глубиной  $U_0$ . Внутри потенциальной ямы электроны движутся свободно, мы пренебрегаем их взаимодействием и будем считать, что они образуют идеальный электронный газ. Поскольку пространственные размеры потенциальной ямы (размеры образца металла) велики по сравнению с межатомными расстояниями, уровни энергии системы образуют почти непрерывный

---

<sup>1</sup>Имя Арнольда Зоммерфельда не слишком хорошо известно широкой публике. Он первым (1904) подумал об излучении заряда, движущегося равномерно, но быстрее скорости света, о чем вспомнили лишь после открытия черенковского излучения (см. примечание 4 на с. 198). Его обобщение «старой» квантовой теории Бора на случай эллиптических орбит электронов с учетом релятивистских эффектов было «перекрыто» более фундаментальными работами в этой области. Полуквантовая теория свободных электронов уступила место гораздо более сложной зонной теории металлов. Зоммерфельд стоял у истоков квантовой теории и был автором первого учебника по квантовой механике, который неоднократно совершенствовался и обновлялся, так что этот монументальный двухтомник не теряет своего значения до сих пор. Им написан также замечательный пятитомный курс теоретической физики. Зоммерфельд создал целую школу теоретической физики, воспитанниками которой были пять нобелевских лауреатов по физике и химии (В. Гейзенберг, П. Дебай, В. Паули, Л. Полинг, Х. Бете). Он сам этой престижной награды не был удостоен (как утверждают, его неоднократно номинировали, но всякий раз очень уж сильны были коллеги-соперники). Тем не менее, А. Зоммерфельда по праву можно причислить к ряду великих физиков XX в.





**Рис. 41.6.** Размещение электронов металла по уровням энергетической зоны (потенциальной ямы):  $U_0$  — глубина потенциальной ямы,  $E_F$  — максимальная (при  $T = 0$ ) энергия электронов, а  $A = U_0 - E_F$  — работа выхода; при нулевой температуре (а) на каждом уровне помещается по два электрона с противоположно направленными спинами; при отличной от нуля температуре (б) часть электронов с верхних заполненных уровней переходит на свободные уровни с энергией  $E > E_F$  — тепловое возбуждение затрагивает небольшую долю электронов в полосе энергий порядка  $k_B T$  вблизи  $E_F$

спектр. Как уже отмечалось в предыдущем разделе, на каждом уровне, согласно принципу Паули, может одновременно находиться не более двух электронов с противоположными ориентациями спинов. Таким образом, валентные электроны металла будут попарно заполнять энергетические уровни, начиная с дна зоны, которое мы примем за начало отсчета энергии. Если полное число электронов в металле первой группы равно  $N$ , то *при абсолютном нуле температуры* будут заполнены первые  $N/2$  состояний с энергиями от нуля до некоторого максимального значения, которое мы обозначим через  $E_F$  (смысл этой величины, которая получила название *энергии Ферми*, будет ясен при дальнейшем рассмотрении). Все остальные состояния с  $E \geq E_F$  будут свободны от электронов.

Полная энергия электрона в металле складывается из потенциальной и кинетической энергий. Схематически энергетические уровни зоны проводимости показаны на рис. 41.6, где  $U_0$  — *глубина потенциальной ямы*, характеризующая энергию, которую необходимо сообщить электрону, находящемуся на самом нижнем уровне зоны проводимости, чтобы удалить его за пределы металла. Величина  $A$  определяет *работу выхода*, т. е. наименьшую энергию, которую необходимо сообщить электрону для того, чтобы удалить его из твердого тела в вакуум. В соответствии с вышесказанным работа выхода определяется соотношением  $A = U_0 - E_F$ . При  $T = 0$  (см. рис. 41.6, а) число электронов на всех заполненных уровнях равно двум, для незаполненных оно равно нулю. Подчеркнем, что согласно принципу запрета Паули электроны вынуждены попадать в возбужденные энергетические состояния даже при абсолютном нуле температуры — согласно классической физике все электроны должны были бы образовать «конденсат» с нулевой энергией.

При отличной от нуля температуре часть электронов переходит на свободные уровни с энергией  $E > E_F$ . Поскольку энергия теплового возбуждения равна по порядку величины  $k_B T$ , при относительно невысоких температурах  $T \ll E_F/k_B$  тепловое возбуждение затрагивает лишь небольшую долю электронов, как это видно из рис. 41.6, б.

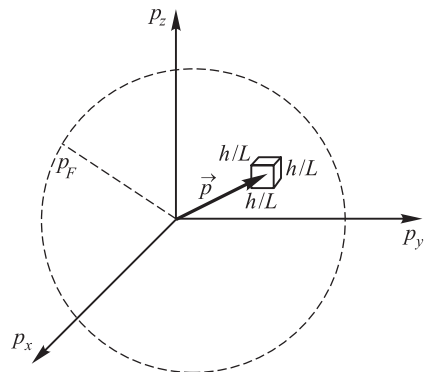
### Электронный газ при нулевой температуре

В металле энергия  $E_F$  наивысшего занятого уровня (при  $T = 0$ ) расположена ниже уровня энергии вакуума и поэтому отрицательна по отношению к нему (см. рис. 41.6). Но если за начало отсчета энергий взять дно валентной зоны, то энергия  $E_F$  валентных электронов металла, обладающих наибольшим импульсом при  $T = 0$ , положительна. Найдем *энергию самого высокого из заполненных энергетических состояний электронов металла при абсолютном нуле* температуры, т. е. *энергию Ферми*  $E_F$ .

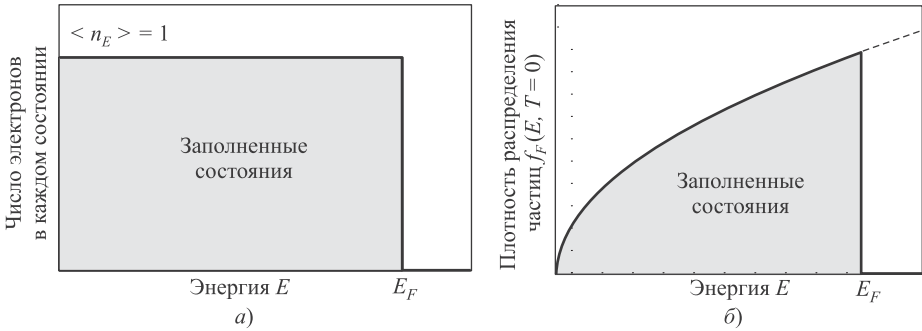
В импульсном пространстве импульсы электронов, обладающих максимальной энергией, лежат на сферической поверхности, имеющей максимальный радиус  $p_F = \sqrt{2mE_F}$  (рис. 41.7). Эта поверхность называется *сферой Ферми*. Импульсы остальных электронов, заполняющих зону проводимости, соответствуют радиусам-векторам, лежащим внутри сферы Ферми, а число состояний  $N$  равно полному числу электронов  $N$  в зоне проводимости, коль скоро в одном состоянии может находиться лишь один электрон (имеется в виду, что состояния, отличающиеся лишь ориентацией спинов электронов, считаются различными). Тогда полное число электронов можно найти, интегрируя (40.57):

$$N = 4\pi V \frac{(2m_e)^{3/2}}{(2\pi\hbar)^3} \int_0^{E_F} \sqrt{E} dE = \frac{8\pi}{3} V \frac{(2m_e E_F)^{3/2}}{(2\pi\hbar)^3}. \quad (41.3)$$

Здесь мы уже полагали, что  $g = 2s + 1 = 2$ , так как для электронов спин  $s = 1/2$ . Вводя концентрацию  $n = N/V$  электронов, получаем из (41.3) формулу



**Рис. 41.7.** Элементарная ячейка в импульсном пространстве имеет линейный размер  $h/L$ . Весь коллектив электронов при  $T = 0$  располагается внутри сферы радиусом  $p_F$



**Рис. 41.8.** Число ферми-частиц в каждом состоянии (а) и плотность их распределения по энергиям (б) при  $T = 0$

для энергии Ферми

$$E_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}. \quad (41.4)$$

При нулевой температуре  $T = 0$  электроны последовательно занимают состояния, начиная с самого низшего, так что на каждое состояние (с учетом спина) приходится один электрон. Иными словами, **числа заполнения**  $n_k$  состояний с номером  $k$  (с учетом спина) могут быть равны лишь нулю или единице. Значит, число электронов  $\langle n_E \rangle$  в каждом состоянии, энергия которого меньше энергии Ферми, равно единице, а в состояниях с  $E > E_F$  — нулю:

$$\langle n_E \rangle_{T=0} = \begin{cases} 1, & E \leq E_F, \\ 0, & E > E_F. \end{cases} \quad (41.5)$$

График числа частиц в состоянии с энергией  $E$  имеет вид ступеньки, показанной на рисунке 41.8, а. Умножая  $\langle n_E \rangle_{T=0}$  на число состояний  $dN$  (см. (40.57)), получаем распределение  $dN = f_F(E, T = 0) V dE$ , где **функция (плотность) распределения**  $f_F(E, T = 0)$  **ферми-частиц по энергиям при нуле температуры** равна:

$$f_F(E, T = 0) = 4\pi V \frac{(2m_e)^{3/2}}{(2\pi\hbar)^3} \begin{cases} \sqrt{E} dE, & E \leq E_F, \\ 0, & E \geq E_F. \end{cases} \quad (41.6)$$

Эта функция показана на рисунке 41.8, б. На обоих рисунках заштрихована область  $E \leq E_F$  внутри сферы Ферми, заполненная электронами.

**Задача 41.63.** Найти среднюю энергию  $\langle E \rangle$  электрона в металле при  $T = 0$ .

**Решение.** Вероятность того, что электрон имеет энергию в интервале  $(E, E + dE)$ , равна  $dW = dN/N$ . Разделив (40.57) на (41.3), найдем:

$$dW = \frac{3}{2E_F^{3/2}} \sqrt{E} dE. \quad (41.7)$$

Теперь среднюю энергию электрона можно найти простым интегрированием выражения  $E dW$  по энергиям, которыми обладают все электроны данной зоны при  $T = 0$ , т. е.

от 0 до  $E_F$ :

$$\langle E \rangle = \int_0^{E_F} E dW = \frac{3}{2E_F^{3/2}} \int_0^{E_F} E^{3/2} dE = \frac{3}{2E_F^{3/2}} \left( \frac{2}{5} E_F^{5/2} \right) = \frac{3}{5} E_F. \quad \blacksquare \quad (41.8)$$

Приведем численные оценки. Если использовать для среднего значения концентрации электронов проводимости в металлах  $n \sim 5 \cdot 10^{28} \text{ м}^{-3}$ , получим  $E_F = 5 \text{ эВ}$ ,  $\langle E \rangle = 3 \text{ эВ}$  и **температуру Ферми** (мы еще встретимся с этим важным параметром)  $T_F = E_F/k_B = 6 \cdot 10^4 \text{ К}$ . Найдем также максимальную и среднюю скорости электронов при абсолютном нуле:  $v_F = \sqrt{2E_F/m_e} = 1,4 \cdot 10^6 \text{ м/с}$  и  $\langle v \rangle = \sqrt{2\langle E \rangle/m_e} = 1,1 \cdot 10^6 \text{ м/с}$ . Оказывается, скорости электронов очень велики даже при абсолютном нуле температуры. Максимальная энергия электронов  $E_F$  меньше глубины потенциальной ямы  $U_0$  (см. рис. 41.6). Поэтому, несмотря на большие скорости, электроны вылететь из металла при  $T = 0$  не могут.

Итак, валентные электроны в металле при  $T = 0$  не находятся в состоянии покоя, как это следовало бы ожидать, исходя из классических представлений, а движутся с различными скоростями. Средняя скорость этого движения весьма велика. Несмотря на это, электронный газ вносит очень малый вклад в теплоемкость металлов. Действительно, средняя тепловая энергия  $k_B T$  при комнатной температуре составляет примерно  $1/40 \text{ эВ} \ll E_F$ . Такая энергия может возбудить только электроны, находящиеся на самых верхних энергетических уровнях, примыкающих к уровню Ферми. Основная же масса электронов, размещенных на более глубоких нижних уровнях, остается в прежних состояниях, т.е. они не поглощают энергию и не нагреваются. Таким образом, поскольку на практике всегда  $T \ll T_F$ , в процессе нагревания металла участвует лишь незначительная часть электронов проводимости, доля которых примерно равна  $T/T_F$ . Соответственно отношение «квантовой» теплоемкости к классической (предполагающей, что в тепловом движении участвуют все электроны) имеет при комнатной температуре порядок  $c_{эл}/c_{реш} \sim T/T_F \sim 300/6 \cdot 10^4 = 5 \cdot 10^{-3}$ . Другими словами, при комнатной температуре вклад электронного газа в полную теплоемкость примерно в 200 раз меньше вклада решетки!

Из молекулярной физики известно (см. т. 1, разд. 10.3), что для одного моля идеального газа

$$pV = RT = \frac{2}{3} N_A \left( \frac{3}{2} \frac{R}{N_A} T \right) = \frac{2}{3} N_A \left( \frac{3}{2} k_B T \right) = \frac{2}{3} N_A \langle E \rangle,$$

т.е.  $p = \frac{2}{3} n \langle E \rangle$ . Подставляя  $n = 5 \cdot 10^{28} \text{ м}^{-3}$  и  $\langle E \rangle \sim 3 \text{ эВ} = 4,8 \cdot 10^{-19} \text{ Дж}$ , получаем оценку  $p \sim 2 \cdot 10^{10} \text{ Па} \sim 2 \cdot 10^5 \text{ атм}$ . Такое огромное давление электронного газа в металле дополнительно увеличивается взаимным отталкиванием электронов друг от друга из-за сил электрического взаимодействия. Полное давление электронного газа компенсируется силами притяжения между электронным газом и ионами кристаллической решетки (или «положительным размазанным» зарядом, которым мы заменили совокупность ионов решетки).

## 41.4 Статистика Ферми—Дирака

Снова используем общее выражение (40.49) для вероятности обнаружить  $n_i$  фермионов в состоянии  $i$ :

$$W = A \exp \left( - \sum_i n_i \frac{E_i - \mu}{k_B T} \right).$$

Как и в случае с бозонами, из всего набора возможных состояний системы проследим за каким-то конкретным состоянием  $k$  с энергией  $E_k$ . В нем с вероятностью  $W(0)$  может не оказаться ни одной частицы ( $n_k = 0$ ), а с вероятностью  $W(1)$  — всего одна частица ( $n_k = 1$ ). Из общего выражения (40.49) следует связь этих вероятностей:

$$W(1) = W(0) \exp \left( \frac{\mu - E_k}{k_B T} \right). \quad (41.9)$$

Поскольку третьего не дано, должно выполняться условие равенства единице суммы вероятностей  $W(0) + W(1) = 1$ . Вместе с (41.9) это дает выражения для искомых вероятностей:

$$W(0) = \frac{\exp \left( \frac{E_k - \mu}{k_B T} \right)}{\exp \left( \frac{E_k - \mu}{k_B T} \right) + 1}, \quad W(1) = \frac{1}{\exp \left( \frac{E_k - \mu}{k_B T} \right) + 1}. \quad (41.10)$$

Значит, среднее число ферми-частиц  $\langle n_k \rangle = W(0) \times 0 + W(1) \times 1 = W(1)$  в состоянии  $k$  получается равным

$$\langle n_k \rangle = \frac{1}{\exp \left( \frac{E_k - \mu}{k_B T} \right) + 1}. \quad (41.11)$$

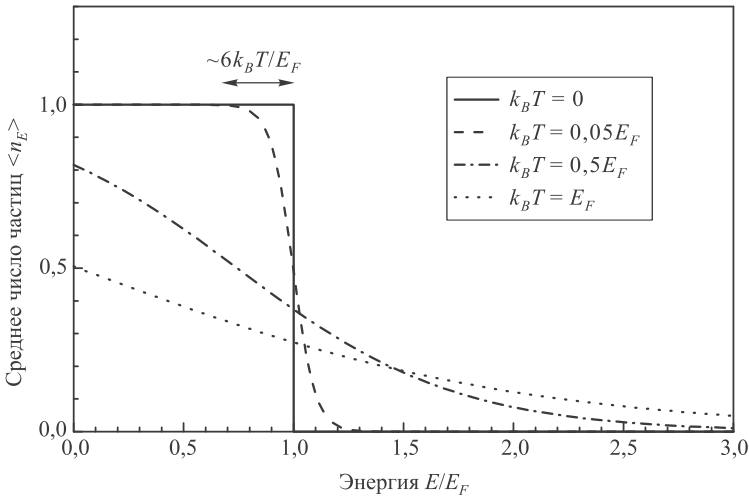
Формула (41.11) является основой *квантовой статистики Ферми—Дирака*, а величина  $\langle n_k \rangle$  называется *функцией распределения Ферми* (или просто *функцией Ферми*). При высоких температурах получаем  $\langle n_k \rangle \approx 1/2$ , т. е. равномерное распределение частиц по состояниям. Если число частиц  $N$  в системе фиксировано, то химический потенциал  $\mu$  определится из условия

$$N = \sum_k \frac{1}{\exp \left( \frac{E_k - \mu}{k_B T} \right) + 1}. \quad (41.12)$$

Подчеркнем также, что согласно (41.11) среднее число частиц в данном состоянии всегда не превышает единицы. Это — прямое следствие принципа Паули.

## 41.5 Распределение Ферми—Дирака

Рассмотрим теперь свойства электронного газа при температурах, отличных от нуля, но еще достаточно низких. Предположим, что температура такова, что



**Рис. 41.9.** Среднее число частиц при разных температурах: «ступенька» при  $T = 0$  (см. уравнение (41.5)) показана сплошной линией; при низких температурах область размытости «ступеньки» вблизи уровня Ферми ( $E/E_F \sim 1$ ) имеет порядок  $\Delta E \approx 6k_B T$

$k_B T$  значительно меньше максимальной энергии электронов  $E_F$ . В таком случае тепловое возбуждение электронного газа будет сравнительно незначительным, так как тепловое возбуждение может переводить электроны из энергетических состояний, заполненных при  $T = 0$ , только в близлежащие более высокие энергетические состояния. Это означает, что тепловое возбуждение недостаточно для того, чтобы поднять электрон с уровня энергии  $E \ll E_F$  на уровень энергии с  $E > E_F$ . Его хватает лишь на возбуждение электронов, находящихся на энергетических уровнях, лежащих вблизи сферы Ферми ( $E \simeq E_F$ ) в узком интервале энергий  $\sim k_B T$ . Часть электронов с этих уровней оказывается переброшенной на уровни, лежащие выше уровня  $E = E_F$ , но отстоящие от него не дальше чем на  $k_B T$ . На рисунке 41.6, б схематически изображено тепловое возбуждение при низких температурах ( $T \ll E_F/k_B$ ).

Часть уровней, лежащих ниже  $E_F$ , оказывается освобожденной чаще всего от одного из заполняющих их электронов. На уровнях же, лежащих выше  $E_F$ , появляются одиночные электроны. Функция распределения Ферми изменяется. Если при  $T = 0$  она представлялась ступенчатой кривой (см. рис. 41.8, а), то при отличных от нуля температурах она принимает вид, изображенный на рис. 41.9. Понижение кривой при  $E < E_F$  с ростом температуры означает, что среднее число электронов в данном состоянии оказывается меньшим единицы — электроны переходят на уровни, лежащие выше  $E_F$ . (Напомним в который раз, что, говоря о состоянии и о среднем числе частиц в нем, мы учитываем направление спина: следует также помнить, что на одном энергетическом уровне могут располагаться два электрона с разными направлениями спинов, т. е. каждому уровню соответствуют два различных состояния.) Область энергий, в которых среднее число электронов в каждом состоянии оказывается меньшим единицы

(но большим нуля, конечно), назовем *зоной размытости*. Из рисунка 41.9 и из сказанного выше следует, что ширина зоны размытости по порядку величины равна нескольким величинам  $k_B T$ . Число электронов, попадающих на уровни, лежащие выше  $E_F$ , весьма мало по сравнению с полным числом электронов. Точно так же число электронов, находящихся на уровне энергии «без пары», составляет малую долю от полного числа электронов.

Для оценки рассмотрим цезий — металл с небольшой энергией Ферми  $E_F = 1,53$  эВ (к примеру, у меди  $E_F = 7,04$  эВ). Значение постоянной Больцмана, где в качестве единицы энергии использован электронвольт, равно  $k_B = 8,617 \cdot 10^{-5}$  эВ/К. Отсюда находим температуру Ферми для цезия:  $T_F = E_F/k_B \approx 18\,000$  К (для меди  $T_F \approx 82\,000$  К. Отсюда следует, что для металлов даже комнатные температуры являются низкими: при  $T = 300$  К в цезии возбуждается небольшая ( $T/T_F = 0,017$ ) доля электронов — менее двух процентов (в меди и того меньше, примерно 0,4%). Это означает, что к таким системам *необходимо* применять квантовую статистику Ферми—Дирака.

Среднее число фермионов (*функция распределения Ферми*)

$$\langle n_E \rangle = \frac{1}{\exp\left(\frac{E - \mu}{k_B T}\right) + 1} \quad (41.13)$$

представляет собой вероятность того, что состояние с энергией  $E$  будет занято электроном. От среднего числа бозонов (см. (40.53)) эта формула отличается лишь знаком перед единицей в знаменателе, но это ведет к кардинальным различиям в поведении ансамблей бозонов и фермионов.

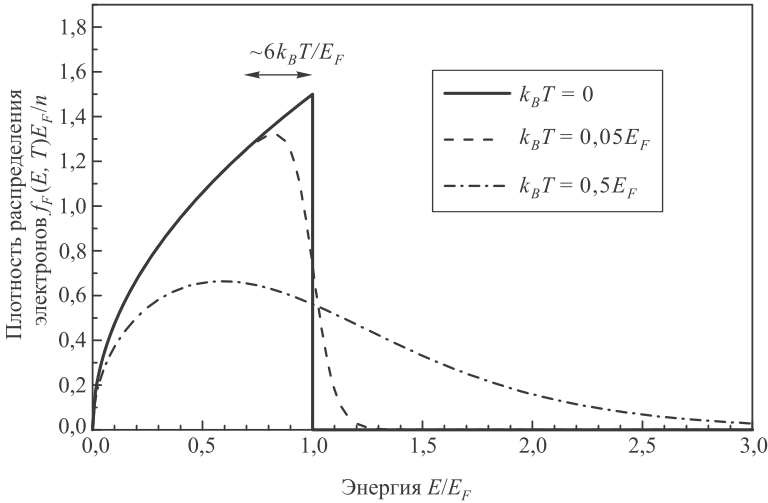
Величина  $\mu$ , как мы уже знаем из раздела 40.6, называется *химическим потенциалом*. В отличие от бозе-частиц, для которых химический потенциал отрицателен, химический потенциал для ферми-частиц при низких температурах должен быть положительным. В противном случае экспонента в знаменателе функции Ферми (41.13) была бы бесконечно велика, а все числа заполнения стремились бы к нулю. Но ферми-частицы не могут выпасть в конденсат: это противоречило бы принципу Паули.

Физический смысл химического потенциала можно понять из анализа поведения функции распределения Ферми вблизи абсолютного нуля. При  $T = 0$  она описывается выражением

$$\langle n_E \rangle = \begin{cases} 1, & E \leq \mu(T = 0), \\ 0, & E \geq \mu(T = 0). \end{cases} \quad (41.14)$$

Сравнивая с (41.5), приходим к выводу, что химический потенциал  $\mu$  при абсолютном нуле равен энергии Ферми:  $\mu(T = 0) = E_F$ . При отличных от нуля температурах химический потенциал представляет собой энергию состояния, вероятность нахождения в котором в точности равна 1/2.

Умножая среднее число частиц (электронов)  $\langle n_E \rangle$  (см. (41.13)) на число состояний  $dN$  (см. (40.57)), получаем, как и в случае бозонов, закон распределения частиц (электронов) по энергиям  $dN = f(E, T)VdE$ , где *функция (плотность) распределения*  $f_F(E, T)$  *электронов, приходящихся в единице объема*



**Рис. 41.10.** Функция распределения электронов по энергиям при разных температурах; ширина зоны размытости имеет тот же порядок, что и на рис. 41.9

на *единичный интервал энергий*, равна:

$$f_F(E, T) = \frac{dN}{VdE} = 4\pi \frac{(2m_e)^{3/2}}{(2\pi\hbar)^3} \frac{\sqrt{E}}{\exp\left(\frac{E-\mu}{k_B T}\right) + 1}. \quad (41.15)$$

Распределение частиц по энергиям с плотностью (см. (41.15)) называется **распределением Ферми—Дирака**.

На рисунке 41.10 изображено поведение функции распределения частиц по энергиям при конечных температурах, причем по оси  $y$  отложены значения безразмерной комбинации  $f_F(E, T)E_F/n$ . Плотность распределения при нулевой температуре показана здесь сплошной линией — это та же кривая, что и на рис. 41.8, б.

Интеграл по всем энергиям дает концентрацию  $n$  электронов в системе:

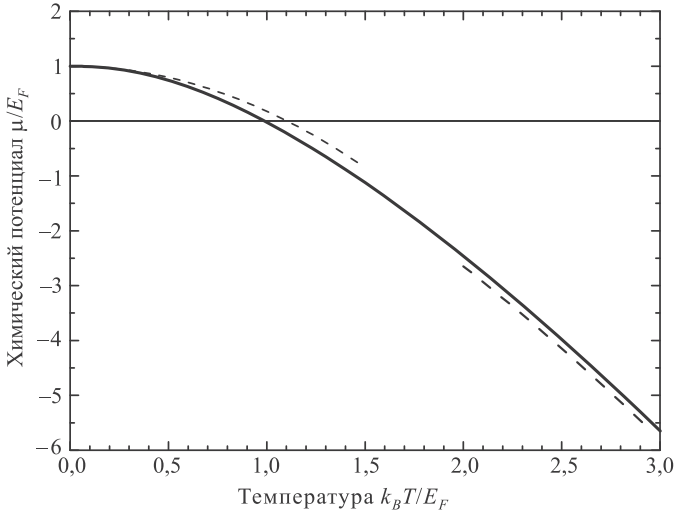
$$n = 4\pi \frac{(2m_e)^{3/2}}{(2\pi\hbar)^3} \int_0^\infty \frac{\sqrt{E}dE}{\exp\left(\frac{E-\mu}{k_B T}\right) + 1}. \quad (41.16)$$

Это соотношение представляет собой уравнение для химического потенциала  $\mu$  как функции температуры.

Оказывается, что химический потенциал хотя и зависит от температуры, но очень слабо. При низких температурах  $k_B T \ll E_F$  он имеет вид

$$\mu \approx E_F \left[ 1 - \frac{\pi^2}{12} \left( \frac{k_B T}{E_F} \right)^2 + \dots \right]. \quad (41.17)$$





**Рис. 41.11.** Зависимость химического потенциала от температуры. При  $T = 0$  химический потенциал в точности совпадает с энергией Ферми. При повышении температуры  $\mu$  убывает и примерно в области  $k_B T \sim E_F$  становится равным нулю, а затем принимает отрицательные значения. На графике пунктирными линиями показаны также приближенные расчеты по формулам (41.17) и (41.18) для низких и высоких температур соответственно

Учитывая малость отношения  $k_B T/E_F$ , во многих случаях (а именно для металлов) можно полагать  $\mu \approx E_F$ .

Значение химического потенциала при высоких температурах можно найти из уравнения (41.16), если пренебречь единицей в знаменателе выражения под интегралом. Тогда интегрирование легко выполняется и мы получаем выражение

$$\mu = -k_B T \left( 0,2847 + \frac{3}{2} \ln \frac{k_B T}{E_F} \right). \quad (41.18)$$

Численный расчет значения химического потенциала по формуле (41.16) представлен на рис. 41.11. Падение химического потенциала с температурой можно заметить также и на рис. 41.9, если вспомнить, что для данной температуры химический потенциал равен энергии, при которой среднее число частиц  $\langle n_E \rangle = 1/2$ . Так, для кривой при  $k_B T = 0,5 E_F$  значение  $\langle n_E \rangle = 1/2$  достигается при  $E = \mu = 0,743 E_F$ . При еще более высоких температурах тепловое возбуждение электронов позволяет им расположиться равномерно по состояниям как с  $E < E_F$ , так и с  $E > E_F$ , и 50%-е заполнение нигде не достигается — химический потенциал имеет отрицательное значение.

С учетом (41.17) получаем при низких температурах поправку к соотношению (41.8) для средней энергии электрона:

$$\langle E \rangle \approx \frac{3}{5} E_F \left[ 1 + \frac{5\pi^2}{12} \left( \frac{k_B T}{E_F} \right)^2 + \dots \right]. \quad (41.19)$$

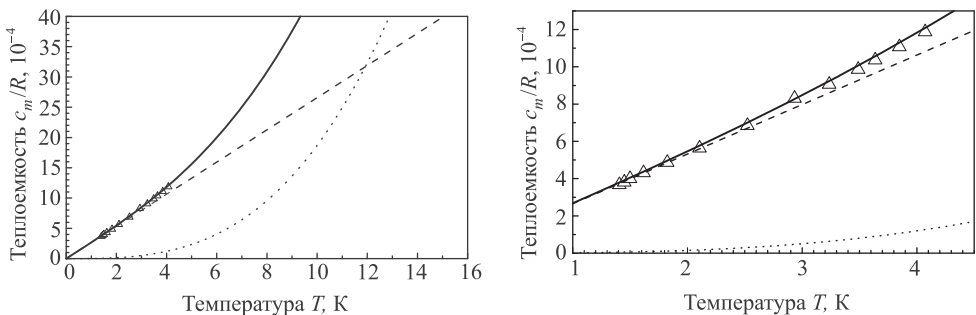
Из этого выражения можно найти теплоемкость электронного газа:

$$C_{эл} = N \frac{\partial \langle E \rangle}{\partial T} = N k_B \frac{\pi^2}{2} \left( \frac{k_B T}{E_F} \right), \quad (41.20)$$

что согласуется с качественными рассуждениями в предыдущем параграфе. Таким образом, теплоемкость электронного газа оказывается линейной функцией температуры и обращается в нуль при  $T = 0$ . Коэффициент пропорциональности содержит только известные величины — универсальные постоянные и число электронов. Для меди, например, теоретическое выражение электронной теплоемкости имеет вид  $C_{эл} = 0,9 \cdot 10^{-4} N k_B T$ , что согласуется с опытными данными с точностью до 2% и при комнатной температуре составляет неизмеримо малую долю теплоемкости кристаллической решетки. Однако теплоемкость решетки зависит от температуры по закону  $C \propto T^3$  согласно (40.32), и при очень низких температурах она становится заметно меньше электронной. На рисунке 41.12 показана расчетная зависимость решеточной, электронной и полной теплоемкостей от температуры вместе с экспериментальными точками. Видно, что сумма решеточной и электронной теплоемкостей отлично описывает эксперимент.

Обсудим теперь зависимость от температуры числа электронов в зоне размытости распределения Ферми. Это число (оно же — число непарных электронов внутри сферы Ферми) называют числом эффективных электронов  $N_{эф}$ . Именно они могут изменять свое состояние под влиянием внешних воздействий, именно эффективные электроны обуславливают теплоемкость электронного газа, его электропроводность. Ясно, что по порядку величины  $N_{эф} \sim T/T_F$ , но сейчас мы хотим найти точное выражение.

Число эффективных электронов можно получить точно так же, как и полное число электронов, с той разницей, что интеграл по энергиям распространяется



**Рис. 41.12.** Температурная зависимость молярной теплоемкости для сплава V (10%)—Cr (90%) при низких температурах: пунктир — вклад кристаллической решетки (температура Дебая  $\theta_D = 500$  К), штриховая линия — вклад электронного газа ( $E_F = 1,6$  эВ), сплошная линия — полная теплоемкость, треугольники — экспериментальные данные; справа показан участок того же графика в увеличенном масштабе

на область энергий  $E \geq E_F$  вне сферы Ферми:

$$N_{\text{эф}} = 4\pi V \frac{(2m_e)^{3/2}}{(2\pi\hbar)^3} \int_{E_F}^{\infty} \frac{\sqrt{E} dE}{\exp\left(\frac{E - \mu}{k_B T}\right) + 1}. \quad (41.21)$$

При низких температурах можно полагать  $\mu = E_F$ . Учитывая также связь (см. (41.3)) энергии Ферми с полным числом частиц и выполняя замену переменных  $E = E_F + xk_B T$ , уравнение (41.21) можно записать в виде

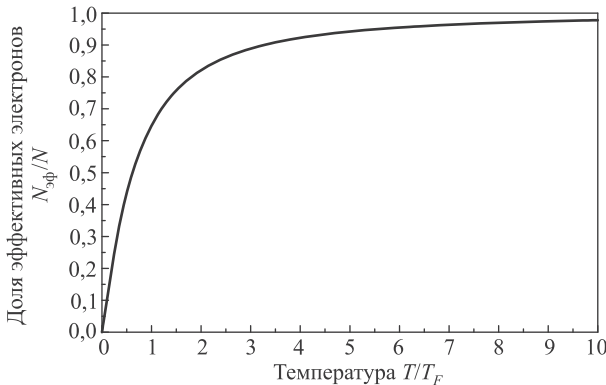
$$N_{\text{эф}} = \frac{3}{2} N \theta \int_0^{\infty} dx \frac{\sqrt{1+x\theta}}{e^x + 1}, \quad \theta = \frac{k_B T}{E_F}. \quad (41.22)$$

При низких температурах квадратный корень в числителе можно заменить на единицу, получающийся интеграл равен  $\ln 2$ , и мы приходим к оценке числа эффективных электронов:

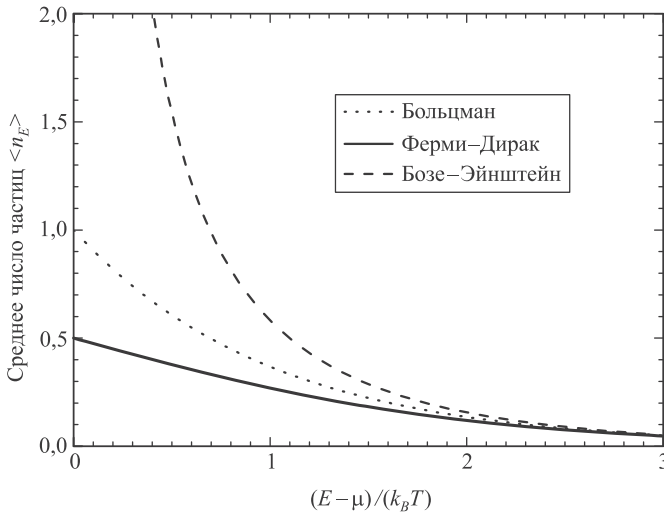
$$N_{\text{эф}} \approx \frac{3 \ln 2}{2} N \frac{k_B T}{E_F} = 1,04 N \frac{k_B T}{E_F} = 1,04 N \frac{T}{T_F}. \quad (41.23)$$

Это тоже подтверждает наши качественные рассуждения. В предыдущем параграфе мы установили, что температура Ферми  $T_F = E_F/k_B$  для металлов составляет несколько десятков тысяч градусов. Поэтому даже при температуре, близкой к температуре плавления металла ( $\sim 10^3$  К),  $N_{\text{эф}}$  составляет малую долю от полного числа электронов в единице объема.

Зависимость доли эффективных электронов  $N_{\text{эф}}/N$  от температуры продемонстрирована на рис. 41.13, где учтено изменение с температурой химического потенциала, показанное на рис. 41.11.



**Рис. 41.13.** Доля эффективных электронов в зависимости от температуры. При  $T = 0$  все частицы находятся внутри сферы Ферми  $E \leq E_F$ ; при  $T = 0,05T_F$  там находится 94,8% частиц; при  $T = 0,5T_F$  — только 56,6% частиц. При очень высоких температурах  $T \gg T_F$  большая часть частиц находится вне сферы Ферми, и к ним применима классическая статистика



**Рис. 41.14.** Сравнение классического распределения Больцмана с квантовыми распределениями Бозе—Эйнштейна и Ферми—Дирака

В заключение проведем анализ того, когда необходимо учитывать законы квантовой статистики, а когда можно применять классическую статистику. Заметим, что если можно пренебречь единицей в знаменателях выражений (40.53) и (41.11), то среднее число электронов в состояниях с различной энергией принимает вид

$$n_k \approx e^{-(E_k - \mu)/(k_B T)} = C e^{-E_k/(k_B T)}, \quad (41.24)$$

т. е. переходит в функцию распределения Больцмана. В этом случае все числа заполнения малы ( $n_k \ll 1$ ) и мы имеем дело с **разреженным ферми-газом**. Такое приближение можно применять, если  $E - \mu \gg k_B T$ . Очевидно, что это неравенство эквивалентно условию  $E \gg E_F$ , т. е. распределение Больцмана справедливо для частиц, находящихся далеко за пределами сферы Ферми. Сделанные утверждения иллюстрируются на рис. 41.14.

Таким образом, поведение электронного газа в сильной степени зависит от соотношения между температурой кристалла и температурой Ферми. Различают два предельных случая:

- $T \ll T_F$ . В этом случае электронный газ называется **вырожденным**. Большая часть электронов находится внутри сферы Ферми, и необходимо учитывать квантовые эффекты.
- $T \gg T_F$ . В этом случае электронный газ называется **невырожденным**. Большая часть электронов находится вне сферы Ферми, и они описываются функцией распределения Больцмана.

Условие вырождения  $k_B T \ll E_F$  может быть записано в виде следующего неравенства:

$$\frac{m_e k_B T}{\hbar^2} \ll 5n^{2/3}, \quad (41.25)$$

т. е. вырождение может быть обусловлено малой массой частиц, или большой плотностью газа, или низкой температурой.

Подставляя в (41.25) характерное для металлов значение концентрации электронов  $n = 5 \cdot 10^{28} \text{ м}^{-3}$ , находим  $T \ll 6000 \text{ К}$ . Иными словами, для электронного газа в металле условие вырождения выполнено вплоть до температур порядка 1000–2000 К. Однако в полупроводниках плотность свободных электронов оказывается много меньшей, чем в металлах. Поэтому уже при комнатной температуре электронный газ во многих полупроводниках является невырожденным и подчиняется классической статистике (описывается распределением Больцмана).

## 41.6 Эффективная масса электрона

Из квантовой механики следует важный результат, что при наличии идеального периодического потенциала электроны проводимости в металле не испытывают никакого сопротивления при движении («окна прозрачности» в оптике). Наблюдаемое в реальных металлах сопротивление возникает вследствие нерегулярностей в периодической структуре, а не из-за взаимодействия электронов с ионными остовами. Нерегулярности, или дефекты, могут быть различного рода, в том числе колебания атомов, вакансии, дислокации, примесные атомы, границы кристаллических зерен или какие-либо другие причины, нарушающие регулярное упорядочение ионных остовов. Это не означает, что регулярная структура кристалла вообще не влияет на движение электронов: влияние существует, и проявляется оно в изменении инертных свойств электронов. Обычно достаточно приписать электрону некоторую *«эффективную массу»*  $m_e^*$  вместо действительной массы  $m_e$ . Электроны тогда будут двигаться через кристалл подобно свободным фермиевским частицам с массой  $m_e^*$ .

В качестве простого примера мы рассмотрим электрон, который может двигаться лишь вдоль одной координаты (квантовая проволочка). Решение волнового уравнения Шрёдингера для свободного электрона в вакууме описывается волновой функцией (волной де Бройля)  $\psi_0(x) = C \exp(ikx)$ , она зависит от волнового числа  $k$ , которое связано с импульсом электрона и его энергией кинематическими соотношениями  $E = p^2/(2m_e) = \hbar^2 k^2/(2m_e)$ . Кристалл мы описываем периодическим потенциалом  $U(x)$ ; если *постоянная решетки* (ее пространственный период) есть  $a$ , то  $U(x+a) = U(x)$ . Как следует из решения уравнения Шрёдингера, волновая функция частицы при таком потенциале описывается выражением  $\psi(x) = u(x) \exp(ikx)$ , похожим на волну де Бройля свободного электрона. Величину  $\hbar k$  называют *квазиимпульсом* (см. разд. 40.5) и соответствующее волновое число  $k$  набирают прямым шрифтом. Амплитуды  $u(x)$  такой волны зависят от координаты  $x$  и тоже периодические:  $u(x) = u(x+a)$ . А вот дисперсионное соотношение (зависимость энергии от квазиимпульса)  $E(k)$  оказывается отличающимся от обычного для свободного электрона квадратичного закона.

Электрон в кристалле можно представить как суперпозицию волн де Бройля с частотами  $\omega$  и длинами волн  $\lambda$ . Максимум амплитуды результирующей волны перемещается в кристалле с групповой скоростью  $u = d\omega/dk$  (см. разд. 31.2), причем наиболее вероятное местонахождение электрона совпадает с пучностями

волновой функции (т. е. с центром группы волн). Групповая скорость  $u$  такого образования представляет собой скорость  $v$  электрона в кристалле. С учетом соотношения  $E = \hbar\omega$  получаем тогда

$$v = \frac{1}{\hbar} \frac{dE}{dk}. \quad (41.26)$$

Напомним, что мы ведем расчеты для одномерной решетки, так что силу, скорость, волновой вектор и т. п. можно рассматривать как скалярные величины.

Рассмотрим поведение электрона под действием приложенного к кристаллу внешнего электрического поля. Действующая на электрон внешняя сила  $F$  за время  $dt$  совершит над электроном работу  $dA = Fvdt$ , что с учетом (41.26) может быть записано как

$$dA = \frac{F}{\hbar} \frac{dE}{dk} dt. \quad (41.27)$$

Увеличение  $dE$  энергии электрона можно выразить как

$$dE = \frac{dE}{dk} dk = \frac{dE}{dk} \frac{dk}{dt} dt. \quad (41.28)$$

По закону сохранения энергии  $dA = dE$ . Приравнявая выражения (41.27) и (41.28), приходим к уравнению изменения квазиимпульса под действием внешней силы

$$F = \hbar \frac{dk}{dt}. \quad (41.29)$$

В свободном пространстве внешняя сила должна быть приравнена по второму закону Ньютона величине  $m_e(dv/dt)$ , а в кристалле мы получили иной закон. Это не означает, конечно, что второй закон Ньютона нарушен: ведь на электрон действуют еще и силы со стороны кристаллической решетки. Однако мы хотим вывести закон движения электрона в кристалле, куда вошли бы лишь внешние силы. Поэтому найдем сначала ускорение электрона:

$$\frac{dv}{dt} = \frac{1}{\hbar} \frac{d}{dt} \left( \frac{dE}{dk} \right) = \frac{1}{\hbar} \frac{d^2 E}{dk^2} \frac{dk}{dt}. \quad (41.30)$$

Используя (41.29), переписываем это уравнение в виде

$$\frac{dv}{dt} = \frac{1}{\hbar^2} \frac{d^2 E}{dk^2} F. \quad (41.31)$$

Теперь, исходя из уравнения (41.31), можно считать, что внешняя сила  $F$  пропорциональна ускорению электрона, а коэффициент пропорциональности и есть **эффективная масса** электрона в кристалле:

$$m_e^* = \frac{\hbar^2}{(d^2 E / dk^2)}. \quad (41.32)$$

В итоге мы получили привычную форму второго закона Ньютона

$$F = m_e^* \frac{dv}{dt}. \quad (41.33)$$

Подчеркнем еще раз, что этот результат не является тривиальным, поскольку ускорение должно быть пропорционально сумме всех сил, приложенных к электрону, — как внешних  $F$ , так и сил  $F_{\text{реш}}$  со стороны ионов, образующих кристаллическую решетку, так что уравнение движения электрона должно иметь вид

$$m_e \frac{dv}{dt} = F + F_{\text{реш}}. \quad (41.34)$$

За устранение сил со стороны кристаллической решетки из правой части уравнения (41.34) пришлось заплатить свою цену — ввести эффективную массу электрона.

Заметим, что в трехмерном случае из-за анизотропии кристалла эффективная масса может иметь разные значения по различным направлениям, т. е. описываться тензором:

$$m_{ij}^* = \frac{\hbar^2}{d^2 E / dk_i dk_j}. \quad (41.35)$$

Чтобы пояснить физический механизм отличия эффективной массы  $m_e^*$  электрона в кристалле от его массы  $m_e$  в вакууме, представим себе движение электрона во внешнем электрическом поле.

Если в некотором (гипотетическом) кристалле атомы расположены далеко друг от друга, то частота перехода электрона от одного атома к другому мала, так как мало перекрытие волновых функций. Во внешнем электрическом поле  $E$  электроны будут чаще переходить от атома к атому в направлении, противоположном полю. Однако особенно ускорить электроны внешним полем не удастся, так как они находятся под действием сильного притягивающего поля своих атомов, поэтому число переходов электронов от одного атома к другому под действием внешнего поля в рассматриваемом случае возрастает незначительно. Поскольку ускорение оказывается небольшим, эффективная масса, фигурирующая в законе Ньютона, должна быть в этом случае больше массы электрона  $m_e$ .

Если атомы сближаются друг с другом, то частота переходов электронов от одного атома к другому увеличивается, и при этих условиях электрическое поле может оказывать на электроны большее ускоряющее действие. Следовательно, эффективная масса электрона уменьшается. По мере увеличения степени перекрытия волновых функций взаимодействие атомов все менее препятствует трансляционному движению электрона. В некоторых материалах электрон может перемещаться внешним электрическим полем быстрее, чем он перемещался бы им в свободном пространстве. Тогда эффективная масса меньше массы электрона. В этом случае движение электронов по атомным орбитам оказывается таким, что оно дополняет переносное действие электрического поля. Таким образом, можно сказать, что в этом случае внутренние атомные силы помогают внешней силе «толкать» электрон в нужном направлении.

Введение эффективной массы позволяет, абстрагируясь от взаимодействия электронов с решеткой, определять характер движения электрона под действием

**Таблица 41.1.** Эффективная масса электрона вблизи дна зоны проводимости в кристаллах некоторых металлов

Металл	Li	Na	K	Ca	Cu	Rb	Cs	Ag	Au
$m_e^*/m_e$	1,19	1,0	0,99	0,83	0,99	0,97	0,98	1,01	1,01

внешнего поля. Приписав электрону массу  $m_e^*$ , можно описать движение электрона под действием силы  $F$ , считая его свободным. Таким образом, все соотношения, полученные в приближении свободных электронов, можно использовать в случае движения электрона в периодическом поле кристаллической решетки, если в них заменить истинную массу  $m_e$  эффективной массой  $m_e^*$ .

В таблице 41.1 приведены эффективные массы электрона вблизи дна зоны проводимости для ряда металлов.

Зависимость энергии электрона от квазиимпульса показана на рис. 41.15, где воспроизведен результат модельного расчета для периодического потенциала притяжения, построенного из «гребенки» дираковских дельта-функций<sup>2</sup>:

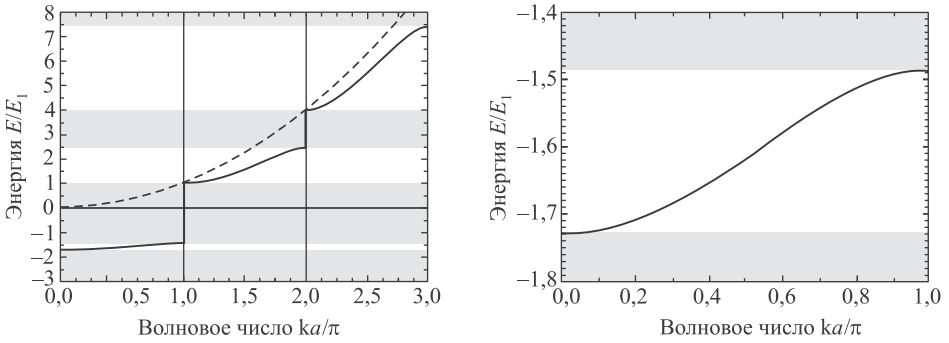
$$U(x) = -g \sum_{n=-\infty}^{\infty} \delta(x - na), \quad (41.36)$$

где  $g$  — константа связи, определяющая силу взаимодействия частицы с нашей одномерной «кристаллической решеткой». Выбор этого потенциала для моделирования решетки обусловлен его простотой: в потенциале притяжения одиночной дельта-функции  $U_1(x) = -g\delta(x)$  имеется всего один уровень, энергия которого отрицательна. В периодической решетке этот уровень расширяется до валентной зоны в области отрицательных энергий. В области положительных энергий, где классическая частица двигалась бы беспрепятственно, также возникают разрешенные и запрещенные (заштрихованы) зоны энергий. Видно, что дисперсионные кривые существенно отличаются от параболической зависимости энергии свободной частицы от импульса (показана пунктиром).

Для наглядности правый график на рис. 41.15 в большем масштабе воспроизводит одну из зон — конкретно нижнюю (валентную). Вблизи дна зон ход кривых  $E(k)$  подобен обычной параболе и эффективная масса положительна. В точке перегиба дисперсионной кривой  $d^2E/dk^2 = 0$ . Следовательно, эффективная масса  $m_e^*$  обращается здесь в бесконечность, т.е. внешнее поле не может оказать никакого воздействия на движение электрона, находящегося в этом энергетическом состоянии. (Здесь энергия зависит от волнового числа по тому же линейному закону, по которому энергия фотона зависит от его импульса, и фотон, как мы знаем, также невозможно ускорить никакими силами.) Вблизи потолка разрешенных зон вторая производная  $d^2E/dk^2 < 0$ , и эффективная масса электронов, занимающих уровни энергии вблизи верхнего края разрешенной зоны, оказывается отрицательной. Это означает, что под совместным действием внешней силы  $F$  и сил  $F_{\text{реш}}$  со стороны решетки электрон получает ускорение,

<sup>2</sup> Дельта-функция Дирака определяется как  $\delta(x) = \begin{cases} +\infty, & x = 0, \\ 0, & x \neq 0, \end{cases}$  и  $\int_{-\infty}^{\infty} \delta(x) dx = 1$ .





**Рис. 41.15.** Модельный расчет зависимости энергии электрона от квазиимпульса в одномерном кристалле: заштрихованы запрещенные зоны энергии, пунктиром показана параболическая зависимость энергии  $E$  от импульса для свободной частицы, на правом графике в более крупном масштабе показана дисперсионная кривая нижней (валентной) зоны модели; энергия дается в единицах энергии  $E_1$  основного состояния в бесконечно высокой потенциальной яме шириной  $a$

направленное в противоположную сторону по отношению к внешней силе  $F$ . Как видно из рис. 41.15, потолку разрешенных зон соответствуют значения волнового числа  $ka = n\pi$ , где  $a$  — период решетки, а  $n$  — натуральное число. Учитывая связь  $k = 2\pi/\lambda$  волнового числа с длиной волны де Бройля, получаем для потолка разрешенных энергетических зон условие  $2a = n\lambda$ . Оно совпадает с **условием Вульфа—Брэгга** при нормальном падении излучения на кристалл (см. разд. 30.6). Таким образом, отрицательная эффективная масса сродни явлению *брэгговского отражения*, когда дифракционный луч лежит по ту же сторону от кристалла, что и падающий. Движение частицы в электрическом поле определяется отношением ее заряда к массе. Смена знака массы эквивалентна смене знака заряда. Иными словами, в таком состоянии электрон будет двигаться так, как двигалась бы положительно заряженная частица с тем же абсолютным значением эффективной массы.

### 41.7 Электропроводность металлов

Классическая теория металлов Друде—Лоренца приводит к следующему выражению для удельного сопротивления металла (см. т. 1, гл. 18, формула (18.64)):

$$\rho = \frac{2m_e v_T}{ne^2 \lambda}.$$

Здесь  $\lambda$  — длина свободного пробега электрона. Согласно классической теории  $\lambda$  не зависит от температуры, так как определяется геометрическими размерами (эффективным сечением  $s$ ) рассеивающих центров (ионов кристаллической решетки) и их концентрацией  $n$ , которую мы считаем совпадающей с концентрацией  $n$  электронов:  $\lambda \sim 1/(ns)$ . Величина  $v_T$  имеет смысл средней скорости теплового движения электронов. В классической теории кинетическая энергия

электронов пропорциональна температуре, поэтому скорость  $v_T$ , а с ней и сопротивление проводника зависят от  $T$  по закону  $\rho \sim v_T \sim \sqrt{T}$ . Такое поведение не согласуется с экспериментом, который в широкой области температур указывает на линейную температурную зависимость удельного сопротивления:  $\rho \sim T$ .

В квантовой теории удельное сопротивление выражается внешне похожей формулой, в которой, однако, следует сделать несколько важных изменений. Во-первых, масса электрона  $m_e$  заменяется на его эффективную массу  $m_e^*$ . Во-вторых, скорость электронов проводимости  $v_T$  необходимо заменить на их фермиеву скорость  $v_F = \sqrt{2E_F/m_e}$ , которая слабо зависит от температуры.

Наконец, в квантовой теории длина свободного пробега электрона может существенно превышать расстояния между ионами. В идеальной периодической решетке электроны вообще не испытывают сопротивления движению, т. е. длина свободного пробега электрона бесконечно велика. Однако строгая периодичность решетки нарушается ее тепловыми колебаниями. Повышение температуры увеличивает амплитуду колебаний ионов, т. е. их эффективный диаметр  $d_{\text{эфф}}$ . Энергия колебаний определяется квадратом амплитуды и пропорциональна температуре, поэтому эффективное сечение рассеивающих центров  $s \sim d_{\text{эфф}}^2 \sim T$ , а длина свободного пробега  $\lambda \sim 1/s \sim 1/T$ . При столкновении с рассеивающими центрами электроны теряют часть своей энергии, что и определяет электрическое сопротивление металла. В квантовой теории этот процесс описывается как рассеяние электронов на фононах — квантах колебаний кристаллической решетки.

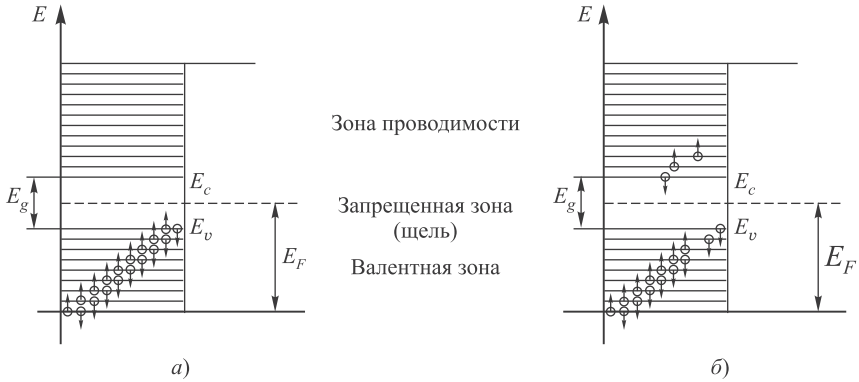
Из сказанного вытекает экспериментально наблюдаемая линейная зависимость удельного сопротивления металлов от температуры:  $\rho \sim T$ . При очень низких температурах эта зависимость сменяется степенным убыванием удельного сопротивления по закону  $\rho \sim T^5$ . Тем не менее даже при очень низких температурах сопротивление проводника стремится не к нулю, а к некоторому конечному значению — *остаточному сопротивлению*, так как электроны рассеиваются и теряют энергию на примесях и дефектах кристаллической решетки.

## 41.8 Полупроводники

Полупроводниками являются вещества, у которых валентная зона полностью заполнена электронами (рис. 41.16, а), а ширина запрещенной зоны невелика (у собственных полупроводников не более 1 эВ). В полупроводниках концентрация подвижных носителей заряда значительно ниже концентрации атомов и может изменяться под влиянием температуры, освещения или относительно малого количества примесей. Эти свойства, а также увеличение проводимости с ростом температуры качественно отличают полупроводники от металлов.

### Дырки

Различают полупроводники с собственной проводимостью (*собственные полупроводники*) и легированные полупроводники с примесной проводимостью (*примесные полупроводники*). В собственном полупроводнике при абсолютном нуле температуры все уровни валентной зоны полностью заполнены электронами, а в зоне проводимости электроны отсутствуют. Электрическое поле не



**Рис. 41.16.** В собственных полупроводниках при  $T = 0$  валентная зона полностью заполнена электронами, а зона проводимости свободна (а). Между ними лежит относительно узкая запрещенная зона (щель). При тепловом возбуждении часть электронов переходит в зону проводимости (б), а на их месте в валентной зоне образуются вакансии (дырки). Уровень Ферми  $E_F$  расположен посередине запрещенной зоны

может перебросить электроны из валентной зоны в зону проводимости, т. е. при  $T = 0$  К собственные полупроводники ведут себя как диэлектрики. (Предполагается, что внешнее электрическое поле меньше пробойного.)

При повышении температуры часть электронов с верхних уровней вблизи потолка ( $E_v$ ) валентной зоны переходит под действием теплового возбуждения на нижние уровни ( $E_c$ ) вблизи дна зоны проводимости (рис. 41.16, б). Таким образом, внешнее электрическое поле получает возможность изменять энергетические состояния электронов, находящихся в зоне проводимости. Кроме того, вследствие образования вакантных уровней в валентной зоне электроны этой зоны также могут изменять свою скорость (энергию) под действием внешнего поля. Следовательно, с ростом температуры электропроводность полупроводника становится отличной от нуля.

Вспомним определение плотности тока  $\vec{j} = -en\vec{v}$  (см. т. 1, гл. 18, формулу (18.7)), где  $n$  — концентрация носителей заряда, а  $\vec{v}$  — их средняя скорость. Здесь мы учли уже, что заряд электрона отрицателен и равен  $-e$ . Среднюю скорость электронов можно записать в виде  $\vec{v} = \frac{1}{N} \sum_i \vec{v}_i$ , где  $N = nV$  — число электронов в объеме  $V$ , а  $\vec{v}_i$  — скорость электрона с номером  $i$ . Подставляя среднюю скорость в выражение для плотности тока, переписываем последнюю в виде  $\vec{j} = -\frac{e}{V} \sum_i \vec{v}_i$ .

В полностью заполненной валентной зоне плотность тока, обусловленная движением электронов, равна нулю, так как для каждого электрона со скоростью  $\vec{v}_i$  найдется электрон с такой же, но противоположно направленной скоростью  $-\vec{v}_i$ . При удалении из валентной зоны электрона, имеющего некоторую скорость  $\vec{v}_m$ , в ней возникает электрический ток:

$$\vec{j} = -\frac{e}{V} \sum_{i \neq m} \vec{v}_i \neq 0. \tag{41.37}$$

Действительно, выделим из суммы скоростей всех электронов скорость  $m$ -го электрона:

$$-\frac{e}{V} \sum_i \vec{v}_i = -\frac{e}{V} \sum_{i \neq m} \vec{v}_i - \frac{e}{V} \vec{v}_m = 0, \quad (41.38)$$

так что (41.37) можно переписать в виде

$$\vec{j} = e\vec{v}_m. \quad (41.39)$$

Из полученного соотношения вытекает, что при отсутствии  $m$ -го электрона в валентной зоне оставшиеся электроны создают ток, эквивалентный току, который создала бы частица с положительным зарядом  $+e$ , движущаяся в том же направлении и с той скоростью, какую имел удаленный электрон. Эта воображаемая частица называется **дыркой**. Таким образом, ток в валентной зоне порождается движением положительного заряда дырки, а движется этот заряд в том же направлении и с той же скоростью, что и удаленный электрон, на месте которого в валентной зоне дырка образовалась.

Вакантные уровни образуются у потолка валентной зоны (см. рис. 41.16), где эффективная масса электрона отрицательна:

$$m_e^* = \frac{\hbar^2}{(d^2 E/dk^2)_v} < 0.$$

Индекс  $v$  служит для напоминания, что вторая производная берется вблизи *потолка валентной зоны*.

Уравнение Ньютона движения электрона во внешнем электрическом поле  $\vec{\mathcal{E}}$  имеет вид  $m_e^* \dot{\vec{v}}_m = -e\vec{\mathcal{E}}$ . Умножив обе части этого уравнения на  $(-1)$ , получаем

$$\left( -\frac{\hbar^2}{(d^2 E/dk^2)_v} \right) \frac{d\vec{v}_m}{dt} = e\vec{\mathcal{E}}. \quad (41.40)$$

Это уравнение описывает движение дырки с положительным зарядом  $+e$  и положительной эффективной массой  $m_h^*$  (индекс  $h$  от англ. *hole* — дырка), равной выражению в скобках в левой части (41.40):

$$m_h^* = -\frac{\hbar^2}{(d^2 E/dk^2)_v}. \quad (41.41)$$

Итак, по своим электрическим свойствам валентная зона с небольшим числом вакантных состояний эквивалентна пустой зоне, содержащей небольшое число дырок. Заметим, что электрон, попавший из валентной зоны в зону проводимости, имеет эффективную массу

$$m_e^* = \frac{\hbar^2}{(d^2 E/dk^2)_c}, \quad (41.42)$$

где индекс  $c$  показывает, что вторая производная энергии по квазиволновому числу берется в районе *дна зоны проводимости*, а не потолка валентной зоны. Из-за разной, вообще говоря, кривизны дисперсионной кривой в этих точках обе массы  $m_e^*$  и  $m_h^*$  положительны, но могут отличаться по величине.

## Собственная проводимость полупроводников

Итак, собственная проводимость возникает в результате перехода электронов с верхних уровней валентной зоны в зону проводимости. При этом в зоне проводимости появляется некоторое число носителей тока — электронов, занимающих уровни вблизи дна зоны. Одновременно в валентной зоне освобождается такое же число мест на верхних уровнях вблизи потолка валентной зоны, в результате чего появляются дырки. Как следствие, полный ток определяется как движением электронов, так и движением дырок.

Обозначив через  $n_e$  и  $n_h$  концентрации электронов и дырок, можно полную плотность тока представить в виде

$$\vec{j} = e(n_h \vec{v}_h - n_e \vec{v}_e), \quad (41.43)$$

где  $\vec{v}_h$  и  $\vec{v}_e$  — дрейфовые скорости дырок и электронов под действием электрического поля.

Найдем положение уровня Ферми для полупроводника. Электроны в зоне проводимости могут рассматриваться как свободные, и для плотности их распределения  $f(E, T)$  мы применим формулу (41.15):

$$f(E, T) = 4\pi V \frac{(2m_e^*)^{3/2}}{(2\pi\hbar)^3} \frac{\sqrt{E - E_c}}{\exp\left(\frac{E - \mu}{k_B T}\right) + 1}. \quad (41.44)$$

Здесь мы, во-первых, заменили массу электрона на его эффективную массу  $m_e^*$  и, во-вторых, использовали под знаком квадратного корня энергию, отсчитываемую от дна зоны проводимости (т. е.  $\sqrt{E - E_c}$  вместо прежнего  $\sqrt{E}$ ), так как по смыслу вывода выражения для  $f(E, T)$  здесь должна стоять кинетическая энергия электронов, а в зоне проводимости последняя как раз и равна  $E - E_c$ . Полное число электронов в зоне проводимости получается интегрированием плотности распределения частиц по всем энергиям, начиная с дна зоны:

$$N_e = 4\pi V \frac{(2m_e^*)^{3/2}}{(2\pi\hbar)^3} \int_{E_c}^{\infty} dE \frac{\sqrt{E - E_c}}{\exp\left(\frac{E - \mu}{k_B T}\right) + 1}. \quad (41.45)$$

Если, как это происходит в полупроводниках при низких температурах,  $E - \mu \gg k_B T$ , то распределение Ферми переходит в распределение Больцмана и можно пренебречь единицей в знаменателе подынтегрального выражения. Делая после этого замену переменной  $x = E - E_c$ , получаем из (41.45) концентрацию электронов  $n_e = N_e/V$  в виде

$$n_e = 4\pi \frac{(2m_e^*)^{3/2}}{(2\pi\hbar)^3} e^{\frac{\mu - E_c}{k_B T}} \int_0^{\infty} dx \sqrt{x} e^{-\frac{x}{k_B T}} = 2 \left( \frac{m_e^* k_B T}{2\pi\hbar^2} \right)^{3/2} e^{-\frac{E_c - \mu}{k_B T}}. \quad (41.46)$$

Для вычисления концентрации дырок в валентной зоне надо произвести следующие изменения в формуле (41.44). Во-первых, нужно учесть, что кинетическая

энергия дырок отсчитывается от потолка валентной зоны (т. е. заменить  $E - E_c$  под знаком квадратного корня на  $E_v - E$ ). Во-вторых, следует учесть и то, что сумма вероятностей обнаружить при данной энергии либо электрон, либо дырку равна единице. Тогда распределение для дырок (при выполнении условия  $\mu - E \gg k_B T$ ) имеет вид

$$1 - \frac{1}{\exp\left(\frac{E-\mu}{k_B T}\right) + 1} = \frac{1}{\exp\left(\frac{-E+\mu}{k_B T}\right) + 1} \approx e^{\frac{E-\mu}{k_B T}}. \quad (41.47)$$

Мы получили то же самое распределение Ферми (или Больцмана) для дырок, но их энергия теперь отсчитывается в направлении отрицательных значений. Соответственно интегрирование должно выполняться по всем возможным энергиям дырок, т. е. от  $-\infty$  до потолка  $E_v$  валентной зоны. Заменяя также эффективную массу электрона на эффективную массу дырки, получаем для числа дырок при низких температурах

$$N_h = 4\pi V \frac{(2m_h^*)^3/2}{(2\pi\hbar)^3} \int_{-\infty}^{E_v} dE \sqrt{E_v - E} e^{\frac{E-\mu}{k_B T}}. \quad (41.48)$$

После замены переменной интегрирования на  $x = E_v - E$  получаем для концентрации дырок

$$n_h = 4\pi \frac{(2m_h^*)^3/2}{(2\pi\hbar)^3} e^{\frac{E_v-\mu}{k_B T}} \int_0^{\infty} dx \sqrt{x} e^{-\frac{x}{k_B T}} = 2 \left( \frac{m_h^* k_B T}{2\pi\hbar^2} \right)^{3/2} e^{-\frac{\mu-E_v}{k_B T}}. \quad (41.49)$$

Перемножая выражения (41.46) и (41.49), получаем важное соотношение

$$n_e n_h = 4 \left( \frac{\sqrt{m_e^* m_h^*} k_B T}{2\pi\hbar^2} \right)^3 e^{-\frac{E_g}{k_B T}}, \quad (41.50)$$

где  $E_g = E_c - E_v$  — ширина запрещенной зоны (*энергетической щели*) между зоной проводимости и валентной зоной. Поскольку при выводе выражений для концентрации электронов и дырок было использовано лишь условие малости температур, соотношение (41.50) применимо не только к собственным полупроводникам. В него входят лишь характеристики вещества (масса, ширина щели), и потому можно сделать вывод, что увеличение концентрации электронов сопровождается уменьшением концентрации дырок.

Однако для собственных полупроводников мы имеем еще *условие электронейтральности кристалла*: полное число электронов в зоне проводимости равно полному числу дырок в валентной зоне. Казалось бы, условие электронейтральности противоречит соотношению (41.50), но вспомним, что в выражения для концентраций  $n_e$  и  $n_h$  входит химический потенциал, причем с разными знаками в экспоненте. Подставляя в равенство  $n_e = n_h$  выражения (41.46) и (41.49), получаем

$$\exp\left(\frac{2\mu}{k_B T}\right) = \left(\frac{m_h^*}{m_e^*}\right)^{3/2} \exp\left(\frac{E_c + E_v}{k_B T}\right). \quad (41.51)$$

Отсюда находим химический потенциал собственного полупроводника при низких температурах:

$$\mu = \frac{E_v + E_c}{2} + \frac{3}{4}k_B T \ln \frac{m_h^*}{m_e^*}. \quad (41.52)$$

Значение химического потенциала при  $T = 0$  дает положение уровня Ферми, который для собственных полупроводников лежит в середине запрещенной зоны (щели):  $E_F = (E_v + E_c)/2$ . При  $T > 0$  химический потенциал совпадает с энергией Ферми, если  $m_h^* = m_e^*$ . При условии  $m_h^* > m_e^*$  он смещается слегка вверх из этого положения, а в противоположном случае  $m_h^* < m_e^*$  он смещается вниз. Энергия  $k_B T$  теплового возбуждения обычно много меньше ширины запрещенной зоны  $E_g = E_c - E_v$ , и смещение химического потенциала от уровня Ферми невелико в масштабе энергии запрещенной зоны, т. е.  $\mu \approx E_F$ .

Итак, *уровень Ферми в собственном полупроводнике лежит в середине запрещенной зоны между валентной зоной и зоной проводимости*. Этот результат имеет для собственных полупроводников фундаментальное значение. При переходе электрона с верхнего уровня валентной зоны на нижний уровень зоны проводимости затрачивается энергия  $E_g$ . В собственном полупроводнике при появлении электрона в зоне проводимости в валентной зоне обязательно возникает дырка, и поэтому можно считать, что затраченная на весь процесс энергия делится на две равные части — половина тратится на образование дырки, а половина — на перевод электрона в зону проводимости. Иными словами, начало отсчета энергии для каждого из этих процессов должно находиться в центре запрещенной зоны, т. е. энергия Ферми представляет такую энергию, от которой происходит возбуждение электронов и дырок. Сказанное подтверждается выражением для распределения электронов (и дырок), которое следует из (41.50) при условии  $n_e = n_h$ :

$$n_e = n_h = 2 \left( \frac{\sqrt{m_e^* m_h^*} k_B T}{2\pi\hbar^2} \right)^{3/2} e^{-\frac{E_g}{2k_B T}}. \quad (41.53)$$

Оно выглядит как распределение Больцмана при энергии возбуждения  $E_g/2$ .

Таким образом, количество электронов, перешедших в зону проводимости, и количество образовавшихся в валентной зоне дырок пропорционально  $e^{-E_g/(2k_B T)}$ , а поскольку электропроводность  $\sigma$  пропорциональна числу носителей тока, то

$$\sigma = \sigma_0 e^{-E_g/(2k_B T)}, \quad (41.54)$$

т. е. электропроводность собственных полупроводников быстро растет с температурой. При нулевой температуре собственная проводимость полупроводников уменьшается до нуля, т. е. они становятся диэлектриками. Соответственно удельное сопротивление полупроводника  $\rho = 1/\sigma$  уменьшается с температурой тоже по экспоненциальному закону.

Логарифм удельного сопротивления  $\ln \rho$  линейно зависит от обратной величины температуры ( $1/T$ ):

$$\ln \rho = \ln \rho_0 + \frac{E_g}{2k_B T}, \quad (41.55)$$

что описывается на графике прямой линией, наклон которой позволяет определить ширину щели  $E_g$ . Эксперименты подтверждают такую зависимость удельного сопротивления полупроводников от температуры.

Наиболее важными собственными полупроводниками являются *кремний* и *германий*, атомы которых принадлежат к IV группе периодической системы элементов Д. И. Менделеева. Они имеют четыре электрона в наполовину заполненной внешней оболочке. В твердом состоянии атомы Si и Ge образуют решетку типа алмаза, в которой четыре внешних электрона связываются ковалентно с четырьмя соседними атомами, в результате чего образуется полностью заполненная валентная зона, т.е. при  $T = 0$  К кремний и германий являются диэлектриками. Однако ширина запрещенной зоны у них невелика: у кремния  $E_g = 1,1$  эВ, у германия  $E_g = 0,75$  эВ. В результате этого уже при сравнительно невысоких температурах значительное число электронов из валентной зоны переходит в зону проводимости, а в валентной зоне одновременно возникают дырки. Дырка начинает мигрировать по кристаллу, как и освободившийся электрон. При встрече электрона с дыркой они могут **рекомбинировать** — электрон займет свободный уровень в валентной зоне, так что *рекомбинация приведет к одновременному исчезновению свободного электрона и дырки*.

При рекомбинации происходит отдача энергии, полученной при образовании электрон-дырочной пары. Один из возможных механизмов отдачи энергии — излучение фотона. Такой механизм (**излучательная рекомбинация**) подобен аннигиляции электрона и его античастицы — позитрона — с испусканием двух квантов электромагнитного излучения. При других механизмах выделяющаяся при рекомбинации энергия может быть непосредственно передана колебаниям кристаллической решетки (**фононная рекомбинация**) либо свободному носителю заряда — электрону или дырке (**ударная рекомбинация**).

Таким образом, в собственном полупроводнике одновременно происходят два процесса: образование электрон-дырочных пар и рекомбинация, приводящая к попарному исчезновению электронов и дырок. Вероятность образования пар быстро увеличивается с температурой. Вероятность процесса рекомбинации пропорциональна как концентрации свободных электронов, так и концентрации дырок. Следовательно, при каждой температуре устанавливается равновесная концентрация электронов и дырок в соответствии с распределением Больцмана (см. (41.53)).

## Примесная проводимость полупроводников

Электрические свойства собственного полупроводника существенно меняются при введении в него примесей атомов другого элемента. Процесс введения примесей называется *легированием*.

**Примесная проводимость** возникает, если некоторые атомы данного полупроводника заменить в узлах кристаллической решетки атомами, валентность которых отличается на единицу от валентности основных атомов.

Пятивалентные атомы, введенные в естественный четырехвалентный полупроводник в качестве примесей, берут на себя четыре ковалентные связи с соседними атомами естественного полупроводника, а пятый электрон пятивалентного

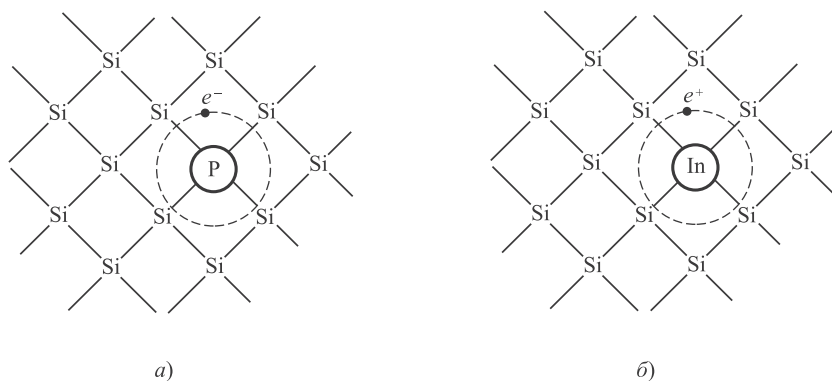


атома оказывается сравнительно слабосвязанным. Этот электрон продолжает принадлежать своему атому, но его энергия связи с ним очень мала, а радиус орбиты велик по сравнению со значениями этих величин для изолированного атома (рис. 41.17).

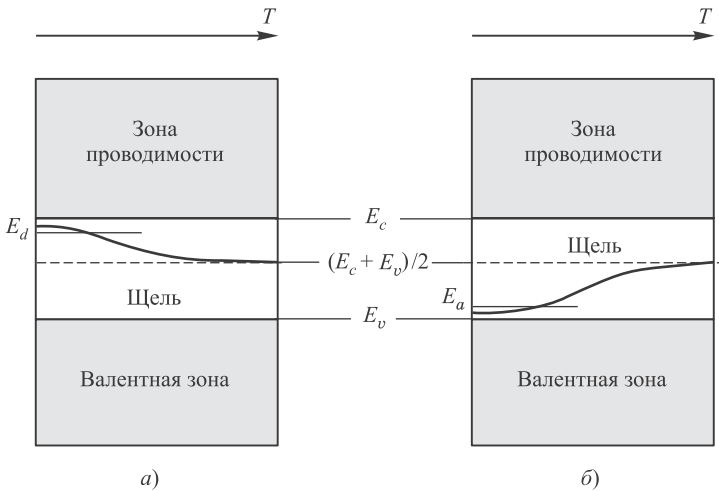
Для оценки энергии связи лишнего электрона можно применить формулу для уровней энергии водородоподобного атома. Следует учесть только, что закон Кулона (взаимодействие электрона с ионом) в среде имеет вид  $-e^2/(4\pi\epsilon_0\epsilon r)$ , где  $\epsilon$  — диэлектрическая проницаемость среды. Соответственно четвертую степень заряда электрона в формуле (37.17) следует заменить на  $e^4/\epsilon^2$ ; кроме того, масса электрона заменяется на его эффективную массу. В результате для основного уровня имеем при  $Z = 1$

$$E_n = -\frac{m_e^* e^4}{2(4\pi\epsilon_0\epsilon)^2 \hbar^2}. \quad (41.56)$$

Типичное значение диэлектрической проницаемости полупроводников  $\epsilon \sim 10$ , а эффективной массы электрона  $m_e^* \sim 0,1m_e$ . В результате получаем, что **энергия ионизации примесного атома** на три порядка меньше энергии ионизации атома водорода (13,6 эВ), т. е. имеет порядок  $E_i \approx 0,01$  эВ. Это означает, что уже при температуре  $T = E_i/k_B \approx 100$  К практически все пентавалентные примесные атомы ионизируются и отдают лишний электрон в зону проводимости. В результате подавляющее число электронов в зоне проводимости при комнатной температуре возникает за счет пятых электронов примесных атомов, а не из-за переходов электронов из валентной зоны. Иными словами, примесная



**Рис. 41.17.** Донорная примесь атомов фосфора (а) и акцепторная примесь атомов индия (б) в кристалле кремния: четыре электрона пентавалентного атома фосфора образуют ковалентную связь, а пятый ( $e^-$ ) оказывается слабосвязанным и может оторваться от атома, начав путешествовать по решетке, т. е. становится электроном проводимости (а); одна из ковалентных связей трехвалентного индия не заполнена — показана как слабосвязанный положительный заряд  $e^+$  около атома примеси, т. е. дырка; если на это место перескочит электрон соседнего атома кремния и процесс будет повторяться, дырка начнет перемещаться по кристаллу



**Рис. 41.18.** Схема энергетических уровней в донорных (а) и акцепторных (б) полупроводниках: при нулевой температуре химический потенциал (уровень Ферми) находится между донорным уровнем  $E_d$  и дном зоны проводимости (а) или между акцепторным уровнем  $E_a$  и потолком валентной зоны (б); при повышении температуры уровень Ферми смещается в центр запрещенной зоны как для чистого полупроводника (показано жирной линией)

электронная проводимость при достаточно низких температурах (вплоть до комнатной) оказывается доминирующей по сравнению с собственной — электронной и дырочной.

Пятивалентные атомы примеси называются **донорными**. Полупроводники с донорной примесью называются также **полупроводниками *n*-типа** (от англ. *negative*), поскольку в них *основными носителями заряда* при достаточно низких температурах (когда собственная проводимость несущественна) являются отрицательно заряженные частицы — электроны.

Энергетические уровни  $E_d$  донорной примеси для слабо легированных полупроводников (полупроводников с малой концентрацией примеси) узки и расположены в запрещенной зоне близко к дну зоны проводимости, поскольку энергия ионизации мала (рис. 41.18, а). Для полупроводников IV группы периодической системы элементов наиболее важными донорными примесями являются элементы V группы: фосфор, мышьяк, сурьма и висмут.

Если в собственный полупроводник IV группы ввести в качестве примеси трехвалентные атомы из III группы элементов (бор, галлий, индий, таллий), то для осуществления ковалентной связи с четырехвалентным окружением этим атомам не хватает по одному электрону. Недостающие электроны они заимствуют у соседних атомов с затратой небольшой энергии порядка 0,01 эВ. В результате в валентной зоне возникают дырки, которые и обуславливают примесную дырочную проводимость полупроводника. Трехвалентные атомы примеси называются **акцепторными**.

Поскольку сродство к электрону атомов акцепторной примеси мало, то энергетические уровни  $E_a$  акцепторной примеси лежат в запрещенной зоне весьма близко к потолку валентной зоны (рис. 41.18, б). Примесные атомы легко присоединяют к себе электроны из валентной зоны, превращаясь в отрицательно заряженные ионы. В валентной зоне при этом образуется дырка. Можно считать, что практически каждый атом примеси порождает дырку в валентной зоне. Полупроводники с акцепторной примесью называются **полупроводниками *p*-типа** (от англ. *positive*), так как основными носителями заряда в них являются положительно заряженные квазичастицы — дырки. Собственная (дырочная и электронная) проводимость при этом играет значительно более слабую роль, чем примесная (при не слишком высоких температурах).

Относительная концентрация примесных атомов чрезвычайно мала, поэтому с ростом температуры концентрация «примесных» носителей тока достигает насыщения, когда практически освобождаются все донорные или заполняются электронами все акцепторные уровни. В то же время с ростом температуры увеличивается собственная проводимость полупроводника, которая при достаточно высокой температуре становится доминирующей. Таким образом, проводимость полупроводника складывается из примесной и собственной проводимостей: при низких температурах преобладает примесная, а при высоких — собственная проводимость.

В заключение этого раздела определим положение химического потенциала в примесных полупроводниках. Уравнения (41.46) и (41.49) остаются в силе, но условие электронейтральности изменяется. Рассмотрим для определенности полупроводник с донорной примесью. Переход электронов с донорного уровня в зону проводимости не приводит к появлению свободных дырок, так как атомы примеси находятся на больших расстояниях друг от друга и вакантное место остается незанятым и не может передвигаться по кристаллу. Если обозначить через  $n_d$  концентрацию атомов примеси (их число в единице объема), то концентрация  $n_d^0$  неионизованных атомов примеси (т. е. концентрация электронов на донорных уровнях) равна, согласно распределению Ферми,

$$n_d^0 = \frac{n_d}{\exp\left(\frac{E_d - \mu}{k_B T}\right) + 1}. \quad (41.57)$$

Концентрация  $n_d^+$  ионизованных атомов примеси определяется как  $n_d^+ = n_d - n_d^0$ , т. е. равна:

$$n_d^+ = \frac{n_d}{\exp\left(\frac{\mu - E_d}{k_B T}\right) + 1}. \quad (41.58)$$

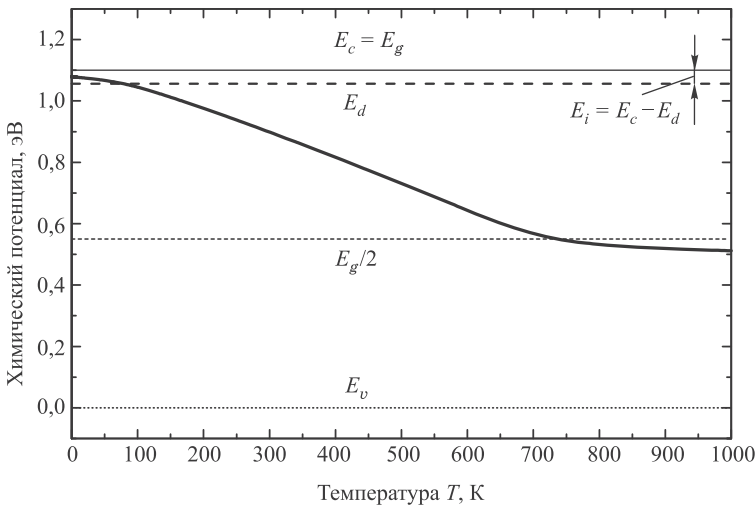
Условие электронейтральности кристалла требует, чтобы число электронов равнялось числу дырок плюс число ионизованных атомов примеси:  $n_e = n_h + n_d^+$ .

Получаем тогда с учетом (41.46) и (41.49):

$$2 \left( \frac{m_e^* k_B T}{2\pi \hbar^2} \right)^{3/2} e^{-\frac{E_c - \mu}{k_B T}} = 2 \left( \frac{m_h^* k_B T}{2\pi \hbar^2} \right)^{3/2} e^{-\frac{\mu - E_v}{k_B T}} + \frac{n_d}{\exp\left(\frac{\mu - E_d}{k_B T}\right) + 1}. \quad (41.59)$$

В качестве примера это сложное трансцендентное уравнение решено численно для значений характеристик системы, близких к кремнию с примесью фосфора (ширина щели  $E_g = 1,10$  эВ, энергия ионизации атома примеси  $E_i = E_c - E_d = 0,044$  эВ, эффективные массы электрона  $m_e^* = 0,98m_e$  и дырки  $m_h^* = 0,52m_e$ , концентрация примеси  $n_d = 10^{22} \text{ м}^{-3}$ ). Результат представлен на рис. 41.19. Ход подобных кривых для разных полупроводников зависит от характеристик вещества и от концентрации примеси. Расчеты показывают, что химический потенциал при низких температурах соответствует середине щели между донорным уровнем и дном зоны проводимости, а при повышении температуры стремится примерно к середине запрещенной зоны как для чистого полупроводника.

Физическая интерпретация такого поведения достаточно ясна. При температурах вблизи абсолютного нуля могут возбуждаться лишь атомы примеси (собственная проводимость полупроводника стремится к нулю). Поскольку уровень Ферми (равный при  $T \sim 0$  химическому потенциалу) соответствует средней энергии возбуждения носителей заряда, он лежит в данном случае посередине



**Рис. 41.19.** Расчетная зависимость от температуры химического потенциала для кремния с примесью фосфора (жирная кривая); потолок валентной зоны взят за начало отсчета энергии ( $E_v = 0$ , пунктир); энергия дна зоны проводимости (тонкая прямая) соответствует ширине щели ( $E_c = E_g$ ); середина щели (тонкая штриховая прямая) выражается  $(E_c + E_v)/2 = E_g/2$ ; донорный уровень  $E_d$  (жирная штриховая прямая) отстоит от дна зоны проводимости на величину энергии ионизации  $E_i$  атома примеси

между донорным уровнем и дном зоны проводимости. При комнатных температурах все атомы примеси уже практически ионизованы, и дальнейшее повышение температуры ведет к росту собственной проводимости, т. е. полупроводник с примесью все больше ведет себя как чистый полупроводник. Соответственно химический потенциал стремится к середине щели с небольшим смещением вниз за счет наличия донорного уровня.

Обобщенное поведение для химического потенциала схематически показано на рис. 41.18, *а* для донорной примеси и на рис. 41.18, *б* для акцепторной.

## Контрольные вопросы

1. Что такое энергетические зоны в кристалле?
2. Объясните механизм формирования энергетических зон при соединении отдельных атомов в кристаллическую решетку.
3. Объясните различие между металлами, диэлектриками и полупроводниками в контексте структуры энергетических зон.
4. Перечислите трудности классической модели свободных электронов в металле.
5. Что такое энергия Ферми? Зависит ли она в металле от объема образца?
6. Что такое температура Ферми? Почему говорят, что электронный газ в металлах вырожден?
7. Объясните физический смысл функции Ферми и ход кривых на рис. 41.9 при разных температурах. При каких условиях распределение Ферми может быть приближенно заменено распределением Больцмана?
8. Что такое плотность состояний? Объясните, почему на рис. 41.10 кривые для  $T = 0$  и  $T = 0,05E_F$  мало отличаются друг от друга.
9. Что такое химический потенциал и какова его зависимость от температуры для системы ферми-частиц?
10. Почему теория теплоемкости Дебая, не учитывающая теплоемкости электронного газа, хорошо работает почти во всем интервале температур? Почему эта теория предсказывает неправильное поведение теплоемкости при очень низких температурах вблизи абсолютного нуля?
11. В чем принципиальное различие статистик Ферми—Дирака и Бозе—Эйнштейна?
12. Что такое эффективная масса электрона? Почему она отличается от массы электрона в вакууме?
13. В чем различие классической и квантовой теорий электропроводности металлов?
14. Как возникают дырки в полупроводниках? Перечислите их свойства.
15. В чем состоит отличие проводимости примесного полупроводника от проводимости собственного? При каком условии их электрические свойства не различаются?
16. Как ведет себя проводимость с ростом температуры?
17. Где находится уровень Ферми в чистом и примесном полупроводниках?
18. Объясните зависимость от температуры химического потенциала полупроводника с донорной примесью (рис. 41.19).

## Глава 42

# Контактные явления

В этой главе мы опишем физические процессы, возникающие при контакте двух проводящих сред с разными свойствами — двух различных металлов или полупроводников. Выше мы посвятили так много времени изучению положения химического потенциала в металлах и полупроводниках, что сразу можно было бы догадаться об исключительно важной роли этой характеристики вещества<sup>1</sup>. Его значение в описании систем с переменным числом частиц предопределено следующим принципом: *при контакте двух разных сред их химические потенциалы выравниваются.*

Электрические явления в контактах разных веществ лежат в основе функционирования многих устройств, так что их изучение важно также и с практической точки зрения.

### 42.1 Работа выхода и контактная разность потенциалов в металле

Носители тока в проводящей среде (в металле — электроны проводимости) находятся в состоянии непрерывного движения, а их наибольшая кинетическая энергия определяется при нуле температуры энергией Ферми  $E_F$ . При отличной от нуля температуре роль энергии Ферми играет химический потенциал  $\mu$ , но для металлов с их огромной температурой Ферми в десятки тысяч градусов отличие  $\mu$  от  $E_F$  невелико даже при температуре в тысячу градусов Цельсия. Наиболее быстрые электроны могут вылететь за пределы металла, который при этом получает избыточный положительный заряд, порождаемый остовом ионов кристаллической решетки. На последующие электроны, пытающиеся вылететь из металла, действуют препятствующие этому силы: притяжение со стороны ионов

---

<sup>1</sup>Химический потенциал — термодинамическая функция состояния, определяющая изменение термодинамических потенциалов, рассмотренных в т. 1, разделе 12.9, при изменении числа частиц в системе. Химический потенциал представляет собой увеличение энергии при добавлении в систему одной частицы без совершения работы. Если полное число частиц в системе не фиксировано, а определяется условиями термодинамического равновесия (как, например, для фотонов в случае теплового излучения АЧТ), то химический потенциал равен нулю.

и этого остова и отталкивание со стороны уже вылетевших электронов вблизи поверхности металла. В результате в пограничной области у поверхности металла образуется *двойной электрический слой*, который можно представить как две близкие друг к другу заряженные поверхности, так что плотности зарядов в противолежащих точках поверхности одинаковы по величине и противоположны по знаку. Толщина этого слоя составляет несколько межатомных расстояний, т. е. исчезающе мала по сравнению с макроскопическими размерами образцов металлов, приводимых в контакт. Вблизи поверхности проводника электрическое поле, создаваемое двойным электрическим слоем, можно представить как поле плоского конденсатора. За пределами двойного слоя поле равно нулю, чему соответствуют постоянные потенциалы  $-\varphi_{\text{внут}}$  для точек внутри металла и  $\varphi_{\text{внеш}}$  для точек вне металла за пределами двойного слоя, но в непосредственной близости к нему. При переходе через двойной слой электростатический потенциал терпит скачок, поэтому для вырывания электрона из металла надо совершить *работу выхода*  $A$  (эта величина уже встречалась нам в разделе 34.2 о фотоэффекте, а также в разделе 41.3 при обсуждении зонной теории металлов). Внутренний потенциал сдвигает энергию всех электронов на величину  $-e\varphi_{\text{внут}}$  (во избежание путаницы скажем, что  $e$  — элементарный электрический заряд, так что заряд электрона  $q_e = -e$ ), поэтому уровень Ферми также сдвигается и становится равным  $E_F - e\varphi_{\text{внут}}$ . Минимальная энергия электрона, вырванного из металла и имеющего нулевую скорость, равна  $-e\varphi_{\text{внеш}}$ . Разность этих энергий и представляет собой работу выхода, откуда находим связь внешнего и внутреннего потенциалов:

$$-e\varphi_{\text{внеш}} = (E_F - e\varphi_{\text{внут}}) + A. \quad (42.1)$$

Мы видим, что внешний потенциал меньше внутреннего потенциала в металле. Когда речь идет об одном металле, можно произвольно выбрать начало отсчета потенциала и положить его равным нулю снаружи металла, так как физический смысл имеет разность потенциалов. Но положение изменяется, если мы рассматриваем контакт двух металлов.

В 1797 г. знаменитый итальянский физик Алессандро Вольта обнаружил, что при соприкосновении двух металлов между ними возникает *контактная разность потенциалов*, зависящая от химических свойств металлов и температуры (*первый закон Вольты*). Проводя эксперименты с различными парами металлов, Вольта выстроил металлы в определенном порядке (в так называемый *ряд Вольты*: Zn, Sn, Pb, Fe, Cu, Pt, Au, Ag, Hg<sup>2</sup>). При контакте двух металлов при одной температуре тот, что стоит в этом ряду впереди, заряжается положительно, а стоящий сзади — отрицательно. Разность потенциалов на концах разомкнутой цепочки, составленной из различных проводников, не зависит от промежуточных металлов в цепочке и определяется лишь контактной разностью потенциалов крайних металлов в цепочке (*второй закон Вольты*).

С точки зрения квантовой теории металлов это явление связано с разным положением уровня Ферми у различных металлов, что, в свою очередь, обуслов-

---

<sup>2</sup>А. Вольта. Новая статья о животном электричестве в трех письмах. Письмо первое. В книге: А. Гальвани и А. Вольта. Избранные работы о животном электричестве. М.-Л.: ОГИЗ, 1937. С. 348.

Таблица 42.1. Работа выхода различных металлов

Металл	Al	Zn	Sn	Pb	Fe	Cu	Ag	Au	Hg
A, эВ	4,25	4,24	4,38	4,0	4,31	4,40	4,3	5,1	4,53

лено разной концентрацией электронов в них. При контакте электроны начинают диффундировать из одного металла в другой, из-за чего они приобретают дополнительный потенциал. Для примера рассмотрим диффузию электронов через контакт меди и цинка. Концентрация электронов в цинке выше концентрации электронов в меди (уровни Ферми этих металлов, как следует из расчета по формуле (41.4), имеют значения  $E_{F,Zn} = 9,37$  эВ и  $E_{F,Cu} = 7,04$  эВ); поэтому электроны из цинка начинают переходить в медь, цинк заряжается положительно, а медь отрицательно — в полном соответствии с рядом Вольты. Под поверхностью контакта металлы приобретают потенциалы  $\varphi_{\text{внут},Cu}$  и  $\varphi_{\text{внут},Zn}$ . Это сдвигает все уровни энергии, так что такой же химический потенциал (в присутствии поля его называют *электрохимическим потенциалом*) становится равным  $\mu - e\varphi$ . Условие совпадения электрохимических потенциалов двух металлов имеет вид

$$\mu_{Cu} - e\varphi_{\text{внут},Cu} = \mu_{Zn} - e\varphi_{\text{внут},Zn}. \quad (42.2)$$

Отсюда находим величину *внутренней контактной разности потенциалов* между точками внутри металлов:

$$\Delta_{\text{внут}}\varphi = \varphi_{\text{внут},Zn} - \varphi_{\text{внут},Cu} = \frac{\mu_{Zn} - \mu_{Cu}}{e}. \quad (42.3)$$

Для металлов даже комнатные температуры можно считать низкими, так что химический потенциал можно положить равным соответствующим энергиям Ферми. Мы видим, что больший потенциал создается в металле, имеющем большую энергию Ферми. Таким образом, внутренняя разность потенциалов равна разности энергий Ферми, отнесенной к заряду электрона. В нашем примере  $\Delta_{\text{внут}}\varphi = 2,33$  В.

Для внешней разности потенциалов имеем уравнение (42.1). Поскольку выражение в скобках в (42.1) одинаково для обоих контактирующих металлов, приходим к выражению для *внешней контактной разности потенциалов*:

$$\Delta_{\text{внеш}}\varphi = \varphi_{\text{внеш},Zn} - \varphi_{\text{внеш},Cu} = -\frac{A_{Zn} - A_{Cu}}{e}. \quad (42.4)$$

Работа выхода для различных металлов приведена в таблице 42.1. Используя эти данные, получаем значение  $\Delta_{\text{внеш}}\varphi = 0,16$  В внешней контактной разности потенциалов для контакта цинка и меди. Эта величина положительна, т. е. внешний потенциал цинка выше внешнего потенциала меди, а это означает, что цинк заряжается положительно, а медь — отрицательно, что соответствует их положению в ряду Вольты.

Из соотношений (42.3) и (42.4) с очевидностью следует второй закон Вольты как для внутренней разности потенциалов, так и для внешней. Пусть, например, проводники 1, 2 и 3 последовательно соединены в цепочку. Имеем для контактов проводников 1-2 и 2-3:

$$\Delta_{\text{внеш}}\varphi_{12} = -\frac{A_1 - A_2}{e}, \quad \Delta_{\text{внеш}}\varphi_{23} = -\frac{A_2 - A_3}{e}. \quad (42.5)$$



Складывая эти уравнения, получаем для внешней контактной разности потенциалов проводников 1 и 3 выражение

$$\Delta_{\text{внеш}}\varphi_{12} + \Delta_{\text{внеш}}\varphi_{23} = -\frac{(A_1 - A_2) + (A_2 - A_3)}{e} = -\frac{A_1 - A_3}{e} = \Delta_{\text{внеш}}\varphi_{13}, \quad (42.6)$$

из которого выпали свойства промежуточного проводника 2. Ясно, что так произойдет при любом числе промежуточных проводников. Если цепь замкнуть, то контактная разность потенциалов будет равна нулю, и ток в такой цепи не пойдет. Это следует из второго закона Вольты при суммировании контактных разностей потенциалов (все равно, внешних или внутренних) при условии, что последний проводник соединен с первым:  $\Delta\varphi_{12} + \Delta\varphi_{23} + \dots + \Delta\varphi_{n-1,n} + \Delta\varphi_{n,1} = \Delta\varphi_{1,1} = 0$ .

Строго говоря, изменение концентрации электронов в контактирующих металлах приводит также к изменению их энергий Ферми (химических потенциалов), что не учитывалось в проведенном анализе. Однако мы имели право так поступить, поскольку изменение концентрации электронов исчезающе мало, что демонстрирует следующая задача.

**Задача 42.64.** Дан медный шарик радиусом  $r = 1$  см. Какое число  $N_e$  электронов надо на него перенести, чтобы потенциал шарика стал равным  $\varphi = 2$  В? Сравнить с числом  $N$  атомов меди в шарике. Плотность меди  $\rho = 8,92$  г/см<sup>3</sup>.

**Решение.** Согласно формуле (16.11) из т. 1, емкость уединенной сферы  $C = 4\pi\epsilon_0 r$ , так что потенциал  $\varphi$  создается на ней зарядом  $q = 4\pi\epsilon_0 r \varphi = 1 \cdot 10^{-2} \times 2 / (9 \cdot 10^9) = 2 \cdot 10^{-12}$  Кл. Разделив  $q$  на элементарный электрический заряд  $|e| = 1,6 \cdot 10^{-19}$ , получим, что такой потенциал создается при переносе всего лишь  $N_e = 1,4 \cdot 10^7$  электронов.

Объем шарика  $V = 4\pi r^3/3 = 4,19 \cdot 10^{-6}$  м<sup>3</sup>, откуда находим его массу:  $m = \rho V = 3,74 \cdot 10^{-2}$  кг. Молярная масса меди  $M_{\text{Cu}} = 63,546$  г/моль (см. таблицу Менделеева). Разделив ее на число Авогадро  $N_A = 6,022 \cdot 10^{23}$  моль<sup>-1</sup> (число частиц в моле вещества), получим массу атома меди  $m_{\text{Cu}} = M_{\text{Cu}}/N_A = 63,546 \cdot 10^{-3} / (6,022 \cdot 10^{23}) = 1,06 \cdot 10^{-25}$  кг. Теперь находим искомое число атомов меди в шарике:  $N = m/M_{\text{Cu}} = 3,5 \cdot 10^{23}$ . ■

Мы видели, что при соединении двух металлических тел (меди и цинка) внутренняя контактная разность потенциалов составляет величину порядка 2 В. Если считать, что каждый атом меди отдает в зону проводимости один электрон, то только что решенная задача позволяет оценить относительное изменение числа (а следовательно, и концентрации) электронов при выравнивании химических потенциалов:  $N_e/N = \Delta n/n = 0,4 \cdot 10^{-16}$ . Так как энергия Ферми  $E_F \propto n^{2/3}$ , на примере меди мы получаем оценку относительного изменения положения уровня Ферми в металле из-за диффузии электронов:  $\delta E_F/E_F = 2\Delta n/(3n) = 3 \cdot 10^{-17}$ . Таким ничтожным сдвигом  $E_F$  действительно можно пренебречь.

## 42.2 Термоэлектрические явления

Гораздо заметнее может оказаться влияние температуры на свойства проводников (в особенности, для полупроводников). В предыдущем разделе рассматривался контакт проводников, находящихся при одной температуре. Однако зависимость химических потенциалов от температуры неизбежно ведет к существованию ряда новых эффектов, известных под общим названием

*термоэлектрических явлений.* В изотропных телах экспериментально наблюдаются три разных термоэлектрических эффекта: термо-ЭДС, эффект Пельтье и эффект Томсона.

## Термо-ЭДС

*Термо-ЭДС* — это напряжение, возникающее на концах проводника (или разомкнутой цепи проводников) вследствие наличия в них градиента температуры. Важно, чтобы концы проводника (или места стыков различных проводников) поддерживались при разной температуре. Цепь, состоящая из двух разных проводников, называется *термопарой*. Возникновение ЭДС в цепи, составленной из разнородных проводников, находящихся при разных температурах, наблюдал в 1821 г. немецкий физик И. Т. Зеебек. Термо-ЭДС возникает из-за перераспределения носителей тока, вызванного градиентом температуры.

Справедливо ради следует отметить, что Зеебек, занимавшийся исследованиями магнетизма, обнаружил отклонение магнитной стрелки вблизи термопары, но отнес это на счет магнитных явлений. Правильное объяснение дал в 1825 г. известный датский физик Х. К. Эрстед, ранее открывший магнитное действие тока. Он и предложил термин *термоэлектричество*. Однако еще в 1794 г. не менее известный итальянский физик А. Вольта сообщил о возникновении тока при нагревании одного из концов проводника. Как известно, в качестве чувствительного электрометра он, следуя Л. А. Гальвани, использовал препарированную лягушку. Вольта наблюдал дергания лягушки под действием электрического тока и понял, что никакого «животного электричества» лягушка не производит, но для возникновения электрического тока концы металлической проволоки, погруженные в два стакана с водой, должны чем-то различаться (и необязательно химическим составом). Вольта проделал эксперимент, когда такое различие создавалось разной температурой на концах дуги. Мы приводим его собственноручное описание этого опыта<sup>3</sup>:

«И вот, согнув дугу из толстой железной проволоки, грубой и гибкой, и погрузив оба ее конца в два стакана с водою, в которую была помещена только что тщательно препарированная лягушка, таким образом, что ее задние лапки находились в одном стакане, а спина или позвоночник (если был оставлен только позвоночник) в другом стакане, мне удавалось заставлять ее сокращаться и подпрыгивать. И, действительно, мне удавалось это лучше всего в первые несколько раз, т. е. два, три, четыре раза, а по прошествии нескольких минут больше не удавалось. Я должен также сказать, что, пользуясь разными железными дугами, я встречал среди них такие, которые с самого начала не давали никакого эффекта. То же самое случалось у меня и с некоторыми серебряными и латунными дугами, которые я находил недействительными при каждой пробе. Таким образом, имеются все основания считать, что с этими негодными дугами ничего не выходило потому, что в обоих своих концах они были совершенно одинаковы по закалке и во всем, а потому равноценны, чего не было у других дуг. Так что здесь как будто обнаруживается трудно и редко попадающееся явление, т. е. настоящее полное сходство во всех отношениях. И вот, перепробовав много дуг и найдя одну из них из железа, которая ничего не давала с самого начала, а в других случаях ожидая, чтобы лягушка ослабела

<sup>3</sup> А. Вольта. Цит. соч. С. 352.

и оказалась уже более не восприимчивою к той из других дуг, которые сначала были в состоянии подействовать на нее (что наступало довольно скоро), я опускал на какие-нибудь полминуты конец такой дуги в кипящую воду, извлекал его оттуда и, не давая ему времени остыть, возвращался к опыту с двумя стаканами холодной воды. И вот тогда-то лягушка в воде и начинала сокращаться, и это происходило даже два, три, четыре раза при повторении опыта до тех пор, пока конец железа, погружавшийся раньше в горячую воду, не остывал вследствие таких более или менее продолжительных повторных погружений или вследствие довольно долгого пребывания на воздухе, и дуга становилась вновь совершенно неспособной возбуждать у животного судороги.»

Выбирая дугу, которая при одинаковой температуре на концах не давала эффекта, Вольта исключил гальванические явления, связанные с контактом металлов с электролитом, и целенаправленно изучал именно воздействие температуры. Таким образом, он бы мог по праву считаться первооткрывателем термоэлектричества, но все же явление возникновения термо-ЭДС часто называется эффектом Зеебека.

Само по себе появление электрического поля в неоднородно нагретом проводнике не ставит в тупик физиков, так как они могут, казалось бы, легко объяснить этот эффект в рамках классической теории металлов. Пусть левый конец металлического проводника поддерживается при температуре  $T_1$ , а правый — при температуре  $T_2 < T_1$ . Более быстрые «горячие» электроны с левого конца будут проникать в правую часть проводника, создавая там избыточный отрицательный заряд, а на левом, более нагретом конце возникает избыточный положительный заряд. Потенциал левого конца металлического проводника окажется выше потенциала правого конца. Иными словами, в неравномерно нагретом проводнике создается электрическое поле, направленное против градиента температуры. Если носители тока заряжены положительно (дырки в полупроводниках), то направление электрического поля совпадает с градиентом температуры. Бесконечно малое изменение температуры должно привести к бесконечно малому изменению потенциала:

$$d\varphi = \alpha dT, \quad (42.7)$$

где коэффициент пропорциональности  $\alpha$  называется *абсолютным коэффициентом термо-ЭДС* материала. Его легко вычислить в рамках классической теории. Действительно, при равновесии в каждом сечении проводника средняя энергия электронов должна быть одной и той же. В классической физике на каждую степень свободы приходится одна и та же кинетическая энергия, равная  $3k_B T/2$ , и для энергии электронов должно соблюдаться равенство

$$\frac{3}{2}k_B T - e\varphi = \text{const.} \quad (42.8)$$

Дифференцируя по  $T$ , получаем оценку коэффициента термо-ЭДС в рамках классической теории:

$$\alpha_{\text{кл}} = \frac{3k_B}{2e}. \quad (42.9)$$

Отсюда вытекает оценка разности потенциалов на концах металлического проводника, поддерживающихся при разной температуре  $T_1$  и  $T_2$ : для конечного

изменения температуры вдоль проводника имеем

$$\Delta\varphi = \varphi_1 - \varphi_2 = \alpha_{\text{кл}}(T_1 - T_2). \quad (42.10)$$

Именно эта разность потенциалов, возникающая на концах неравномерно нагретого проводника, называется *термоэлектродвижущей силой (термо-ЭДС)*. Используя значение постоянной Больцмана, где энергия дана в электронвольтах ( $k_B = 8,617 \cdot 10^{-5}$  эВ/К), приходим к численной оценке  $\alpha_{\text{кл}} = 0,129$  мВ/К.

Эксперимент дает для металлов заметно меньшую величину. К тому же в классической теории этот коэффициент одинаков для всех металлов и не зависит от температуры. И то и другое не соответствует экспериментальным данным.

Дело несколько подправила полуквантовая теория свободных электронов в металле, где точно так же потенциал можно связать с кинетической энергией электрона, только последняя теперь выражается формулой (41.19). Тогда условие равновесия в металле выглядит следующий образом:

$$\frac{3}{5} E_F \left[ 1 + \frac{5\pi^2}{12} \left( \frac{k_B T}{E_F} \right)^2 + \dots \right] - e\varphi = \text{const}. \quad (42.11)$$

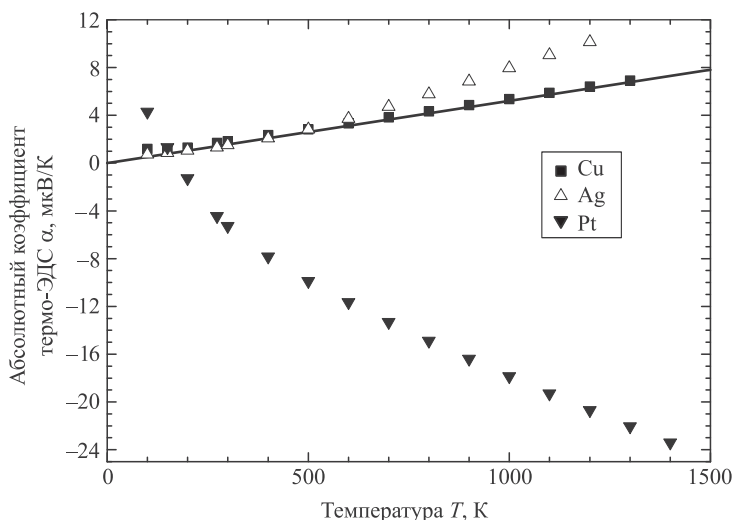
Отсюда получаем, что в аналогичном соотношении  $d\varphi = \alpha dT$  между градиентами потенциала и температуры абсолютный коэффициент термо-ЭДС определяется теперь иным выражением

$$\alpha \approx \pi^2 \frac{k_B}{2e} \frac{T}{T_F} = 0,425 \frac{T}{T_F} \text{ мВ/К}. \quad (42.12)$$

Например, для меди температура Ферми  $T_{F,\text{Cu}} = 81\,600$  К, и при  $0^\circ\text{C}$  ( $T = 273$  К) получаем  $\alpha = 0,475$  мкВ/К, что в 90 раз меньше результата, даваемого классической теорией. Но главные отличия этой оценки от классического результата в другом. Во-первых, через значение уровня (температуры) Ферми появилась зависимость коэффициента термо-ЭДС от материала проводника. Во-вторых, сам коэффициент термо-ЭДС оказался зависящим от температуры проводника. Поэтому линейное соотношение (42.10) между термо-ЭДС и разностью температур на концах проводника справедливо лишь для небольших перепадов температур  $T_1 - T_2 \ll T_2$ . На рисунке 42.1 черными квадратиками показаны экспериментальные данные<sup>4</sup> о температурной зависимости абсолютного коэффициента термо-ЭДС для меди, а тонкой прямой — оценка (42.12). Видно хорошее совпадение нашей модели с экспериментом. Но уже данные по серебру (светлые треугольники) демонстрируют отклонения от линейного закона, а для платины (черные треугольники) абсолютный коэффициент термо-ЭДС и вовсе убывает с ростом температуры и даже меняет знак. Все это показывает, что теория свободных электронов недостаточна для описания термоэлектрических явлений, хотя и ухватывает какие-то их черты. Это же можно сказать и о теории теплоемкости и проводимости металлов. Здесь следовало бы применить зонную теорию металлов и учесть сложную структуру ферми-поверхностей в разных кристаллах. Но эта теория слишком трудна для изложения в курсе общей физики и к тому же было бы слишком опростетчиво утверждать, что термоэлектрические явления понятия полностью и окончательно.

Как мы видели, термо-ЭДС в металлах очень мала. Более отчетливое проявление термоэлектрических свойств следует ожидать в полупроводниках. В

<sup>4</sup>Здесь и далее в этой главе экспериментальные данные взяты из книги «Физические величины. Справочник». М.: Энергоатомиздат, 1991. С. 559–567.

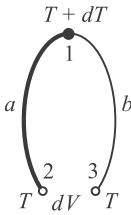


**Рис. 42.1.** Экспериментальные данные о зависимости абсолютного коэффициента термо-ЭДС от температуры для некоторых металлов и теоретическая прямая, наклон которой вычислен для меди по формуле (42.12): видно хорошее согласие с экспериментом

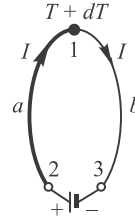
металлах электронный газ вырожден, концентрация носителей тока очень велика и слабо (как и их энергия) зависит от температуры. Этим объясняются малые значения термо-ЭДС. В полупроводниках  $n$ -типа концентрация носителей тока не столь велика и сильно зависит от температуры, а в остальном механизм возникновения термо-ЭДС мало отличается от описанного выше. В полупроводниках  $p$ -типа происходит диффузия дырок. Поскольку они имеют положительный заряд, горячий конец проводника зарядится отрицательно, и его потенциал понизится, а холодный конец зарядится положительно, и его потенциал повысится. В полупроводниках смешанного типа ток переносится и электронами, и дырками. Может даже случиться так, что термо-ЭДС, генерируемые носителями заряда противоположного знака, скомпенсируют друг друга.

## ЭДС термопары

Рассмотрим теперь возникновение термо-ЭДС в термопаре. Если бы термопара состояла из одинаковых материалов, градиенты потенциала, возникающие в проводниках из-за неодинаковой температуры их концов и направленные в разные стороны, компенсировали бы друг друга и результирующая ЭДС была бы равна нулю. Но так как термопары составлены из разных проводников (железо и вода в опытах Вольты, медь и висмут или сурьма в опытах Зеебека), возникающие на их концах разности потенциалов уже не равны друг другу и приводят к появлению так называемой **объемной термо-ЭДС**. Кроме этого, в местах соединения проводников возникает контактная разность внутренних потенциалов. Ранее, оценивая ее величину, мы пренебрегали зависимостью скачка потенциала от



**Рис. 42.2.** К образованию ЭДС в термопаре, составленной из проводников  $a$  и  $b$ . Спай (точка 1) поддерживается при температуре  $T + dT$ , а свободные концы (точки 2 и 3) — при температуре  $T$ ; между свободными концами образуется разность потенциалов  $dV = \varphi_2 - \varphi_3$ , т. е. объемная термо-ЭДС



**Рис. 42.3.** Эффект Пельтье, обратный эффекту возникновения термо-ЭДС: если к термопаре подсоединить источник питания, то текущий по цепи ток  $I$  в зависимости от направления приведет к нагреванию или охлаждению спаю

температуры. Но в термопаре имеются два соединения, и скачок потенциала в них имеет разный знак. Поэтому не зависящие от температуры (и основные) части скачков потенциалов скомпенсируют друг друга, а части, зависящие от температуры, приведут к появлению объемной термо-ЭДС.

Схема возникновения термо-ЭДС показана на рис. 42.2. Рисунок 42.3 иллюстрирует обратный эффект (эффект Пельтье), который будет рассмотрен ниже.

Рассмотрим цепь, составленную из проводников  $a$  и  $b$ , в которой создан бесконечно малый градиент температуры. Пусть горячий контакт находится при температуре  $T + dT$ , холодный — при температуре  $T$ . Цифрой 1 обозначен спай, а цифрами 2 и 3 — свободные концы термопары. Соответственно внутренние потенциалы в указанных точках обозначаем через  $\varphi_i$ ,  $i = 1, 2, 3$ . Используя выражение (42.7) для термо-ЭДС в отдельном проводнике, записываем разности потенциалов как

$$\begin{aligned}\varphi_1 - \varphi_2 &= \alpha_a dT, \\ \varphi_1 - \varphi_3 &= \alpha_b dT.\end{aligned}\tag{42.13}$$

Вычитая первое уравнение из второго, находим разность потенциалов на свободных концах термопары, т. е. ее термо-ЭДС  $dV$ :

$$dV = \varphi_2 - \varphi_3 = (\alpha_b - \alpha_a)dT.\tag{42.14}$$

Применим теперь теорию свободных электронов. Разность  $\alpha_{ba} = \alpha_b - \alpha_a$  называется коэффициентом термо-ЭДС материала  $b$  относительно материала  $a$ . Мы уже получили в (42.12) численную оценку его величины, откуда находим оценку для объемной термо-ЭДС  $dV$  при контакте двух проводников:

$$dV = \alpha_{ba} dT, \quad \alpha_{ba} = 0,425 \left( \frac{1}{T_{F,b}} - \frac{1}{T_{F,a}} \right) T \text{ мВ/К}.\tag{42.15}$$

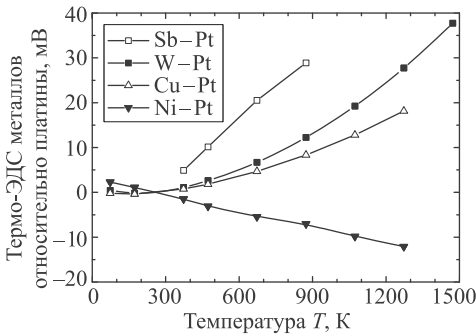
Следует также отметить, что помимо объемной термо-ЭДС существуют и другие механизмы возникновения термоэлектричества: **контактная термо-ЭДС**, появляющаяся вследствие различной зависимости от температуры контактной разности потенциалов проводников в термопаре, и **фононное увлечение**, когда градиент температуры создает поток фононов от горячего контакта к холодному, а те, в свою очередь, увлекают за собой при столкновениях электроны, создавая дополнительную разность потенциалов в термопаре.

До сих пор речь шла о дифференциальных характеристиках термо-ЭДС при бесконечно малых разностях температур. Для конечных разностей, если контакты термопары поддерживаются при температурах  $T$  и  $T_0$ , термо-ЭДС получается интегрированием  $dV/dT$  в пределах от  $T_0$  до  $T$ , а для этого надо знать точную зависимость от температуры коэффициентов термо-ЭДС. Они же зависят от проводников в термопаре, и общетеоретические соображения тут мало помогут — нужны конкретные расчеты для данных материалов. К тому же измеряемые термо-ЭДС относятся всегда к паре проводников и не являются характеристиками отдельного материала. В таблицах обычно приводят термо-ЭДС материалов по отношению к свинцу, платине или меди. В классической модели коэффициент термо-ЭДС не зависит от температуры, и для конечных разностей мы получим линейную зависимость  $V = S_1(T - T_0)$ . В модели свободных электронов коэффициент термо-ЭДС линейно зависит от температуры, и интегрирование  $dV$  дает квадратичную зависимость  $V = S_2(T^2 - T_0^2)$ . Комбинацию этих двух функций издавна, еще со второй половины XIX в., использовали как феноменологическую формулу, аппроксимирующую экспериментальные результаты:

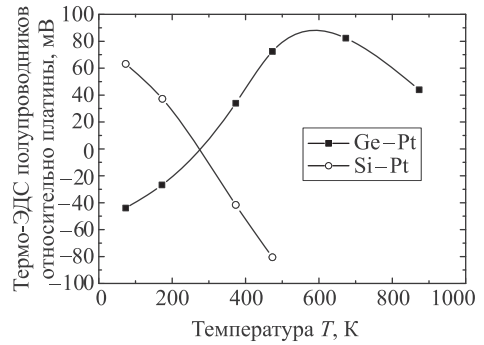
$$V = S_1(T - T_0) + S_2(T^2 - T_0^2). \quad (42.16)$$

Коэффициент  $S_1$  может быть как положительным, так и отрицательным. Для некоторых пар металлов хорошим приближением является линейная аппроксимация ( $S_2 = 0$ ). Этот случай мы видим на графике зависимости термо-ЭДС от температуры для термопары никель—платина (рис. 42.4). Для описания поведения термо-ЭДС в других парах проводников нужна более сложная зависимость. При совпадении температур концов термопары  $T = T_0$  термо-ЭДС обращается в нуль. Но из (42.16) следует, что термо-ЭДС также равна нулю, если при какой-то температуре  $T = T_{\text{ин}}$  окажется, что  $T_{\text{ин}} + T_0 = -S_1/S_2$ . Этот случай можно наблюдать в термопаре вольфрам—платина, где  $S_1 < 0$ ,  $S_2 > 0$ . Если при фиксированной температуре  $T_0$  повышать температуру спаев, то сначала термо-ЭДС положительна и уменьшается с ростом температуры. При температуре  $T_{\text{ин}} = -T_0 - S_1/S_2 \approx 124$  К термо-ЭДС обращается в нуль, при  $T_{\text{мин}} = -S_1/(2S_2) = (T_{\text{ин}} + T_0)/2 \approx 199$  К достигает минимума, при  $T = T_0 = 273$  К снова обращается в нуль и далее возрастает по квадратичному закону. Температура  $T_{\text{ин}}$ , при которой термо-ЭДС проходит через нуль, называется **температурой инверсии**, а температура  $T_{\text{мин}}$  — **температурой нейтральной точки**.

На рисунке 42.5 показаны термо-ЭДС полупроводников (кремния и германия) относительно платины при той же температуре  $T_0 = 273$  К свободных концов термопары, что и на рис. 42.4. Видно, что при одинаковых значениях разности температур  $T - T_0$  термоэлектрический эффект в полупроводниках выражен



**Рис. 42.4.** Экспериментальные данные по температурной зависимости термо-ЭДС некоторых металлов относительно платины при постоянной температуре свободных концов термопары  $T_0 = 273$  К



**Рис. 42.5.** Данные по термо-ЭДС полупроводников при тех же условиях, что и для металлов на рис. 42.4; на обоих графиках положительный знак ЭДС материала соответствует току через спай термопары, идущему к нему от платины

сильнее в несколько десятков раз. Ясно также, что поведение некоторых кривых не может быть описано даже квадратичной функцией (см. (42.16)).

Проводник с большим потенциалом называется положительным (он играет роль анода), а другой — отрицательным (является катодом). Например, при температуре  $T = 473$  К термо-ЭДС никеля равна  $-3,10$  мВ, а меди равна  $+1,83$  мВ. Тогда термо-ЭДС никеля относительно меди при той же температуре равна разности этих термо-ЭДС  $(-3,10 - 1,83)$  мВ =  $-4,93$  мВ, т. е. в термопаре никель — медь никель будет отрицательным электродом (катодом), а медь положительным (анодом). Если вместо никеля взять сурьму (ее термо-ЭДС относительно платины равна  $+10,14$  мВ при  $T = 473$  К), то термо-ЭДС сурьмы относительно меди равна  $(10,14 - 1,83)$  мВ =  $8,31$  мВ, и в этой термопаре проводник из сурьмы будет положительным электродом, а из меди — отрицательным.

## Эффект Пельтье

В 1834 г. француз Ж. Пельтье обнаружил, что при пропускании тока через термопару температура спаев резко меняется и в некоторых случаях наблюдается даже охлаждение.

Наибольшего эффекта Пельтье добился в экспериментах с термопарой висмут–сурьма. Сейчас мы знаем, что при температурах в спаих  $T = 373$  К и  $T_0 = 273$  К коэффициенты термо-ЭДС этих металлов относительно платины равны  $0,90$  мВ для висмута и  $4,89$  мВ для сурьмы, т. е. относительный коэффициент термо-ЭДС сурьмы (анод) относительно висмута (катод) равен  $3,99$  мВ. Пусть в термопаре на рис. 42.2 проводник  $a$  — сурьма, а проводник  $b$  — висмут. Тогда при замыкании цепи ток должен идти через горячий контакт 1 от висмута к сурьме (т. е. против часовой стрелки). Запомним это и обратимся к опытам Пельтье (см. рис. 42.3).



Он спаивал накрест два проводника и пропускал внешний ток через два соседних конца креста. При этом спай либо нагревался, либо охлаждался. Затем Пельтье отключал источник питания и подсоединял гальванометр к двум другим концам креста. Гальванометр показывал наличие термоэлектрического тока (эффект Зеебека), вызванного либо нагреванием, либо охлаждением спая. Современные Пельтье физики не очень доверяли его экспериментам (возможно, из-за его прошлого, так как до 30 лет Пельтье был часовщиком и в открытом им явлении не очень разобрался). В каких случаях спай нагревается, а в каких охлаждается, было установлено лишь позже русским физиком Э. Х. Ленцем (1838). Изучение им эффекта Пельтье показало, что спай нагревался или охлаждался в зависимости от направления внешнего тока. Ленц даже сумел заморозить каплю воды, помещенную в стыке между стержнями из сурьмы и висмута. После отключения внешнего источника направление термоэлектрического тока было обратным направлению внешнего тока, вызвавшего изменение температуры спая. Отсюда следует, что спай 1 на рис. 42.3 должен *нагреваться* при показанном направлении тока от сурьмы к висмуту (по часовой стрелке). В самом деле, если бы спай 1 охлаждался, то термоэлектрический ток после отключения источника питания протекал бы через спай от сурьмы к висмуту, т. е. по-прежнему по часовой стрелке, вопреки данным экспериментов о смене направления тока.

Обратное направление термоэлектрического тока можно понять на основе термодинамического принципа динамического равновесия<sup>5</sup>, сформулированного французским физиком и химиком А. Ле Шателье (1884):

*если система находится в состоянии равновесия, то при внешнем воздействии (изменении температуры, давления, внешнего электромагнитного поля и т. п.) в ней возникают процессы, стремящиеся скомпенсировать оказанное воздействие.*

Частным случаем этого принципа является правило Ленца, определяющее направление индукционного тока. С другим частным случаем мы познакомились только что на примере явления Пельтье: при оговоренных выше условиях спай 1 на рис. 42.3 должен нагреваться, и тогда возникающий термоэлектрический ток будет препятствовать внешнему току  $I$ . При смене направления внешнего тока спай 1 будет охлаждаться.

Теплота, выделяющаяся или поглощающаяся в зависимости от направления тока, проходящего через спай разнородных материалов, называется *теплотой Пельтье*. В отличие от джоулевой теплоты, пропорциональной квадрату тока и никогда не меняющей знак (положительной считаем выделяющуюся теплоту),

---

<sup>5</sup>Принцип динамического равновесия часто называют принципом Ле Шателье—Брауна по имени немецкого ученого К. Ф. Брауна, открывшего его независимо и чуть позже (1887). Но Браун больше известен своими работами в области радио: он обнаружил выпрямляющее действие некоторых кристаллов при прохождении через них переменного тока, изобрел электронно-лучевую трубку, создал детекторный приемник и построил направленную антенну (прообраз современной фазированной антенной решетки). В 1909 г. Браун вместе с итальянцем Г. Маркони получил Нобелевскую премию «за выдающийся вклад в создание беспроволочной телеграфии».

теплота Пельтье  $Q_{\text{П}}$  пропорциональна протекающему через контакт электрическому заряду, т. е. пропорциональна силе тока  $I$  и времени  $t$ :

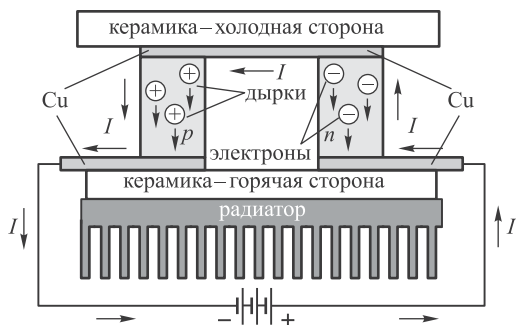
$$Q_{\text{П}} = \Pi_{ab} I t. \quad (42.17)$$

Здесь  $\Pi_{ab}$  — *относительный коэффициент Пельтье*, зависящий от природы материалов  $a$  и  $b$ , находящихся в контакте. Этот коэффициент может быть как положительным, так и отрицательным, в зависимости от того выделяется теплота или поглощается (контакт нагревается или охлаждается). Эффект Пельтье, как и эффект Зеебека, наблюдается лишь при контакте разнородных проводящих материалов.

В металлах эффект Пельтье очень мал и маскируется выделением джоулева тепла и теплопроводностью. Механизм возникновения этого явления связан с затратами или приобретением электронами энергии при прохождении контактной разности потенциалов. Иной механизм выделения теплоты Пельтье наблюдается в контактах полупроводников  $n$ - и  $p$ -типа. Если ток через контакт идет от полупроводника  $n$ -типа к полупроводнику  $p$ -типа, то тепловая энергия тратится на перевод электронов из валентной зоны  $p$ -полупроводника (где образуется свободное место — дырка) в зону проводимости  $n$ -полупроводника. Соответственно этот контакт охлаждается. Если же ток через контакт идет в обратном направлении, от полупроводника  $p$ -типа к полупроводнику  $n$ -типа, то электроны и дырки движутся навстречу друг другу и в области контакта рекомбинируют: электрон уходит из зоны проводимости в  $n$ -полупроводнике и занимает свободное место в валентной зоне  $p$ -полупроводника, уничтожая дырку. В результате рекомбинации выделяется энергия, нагревающая контакт.

Практическое применение нашли *элементы Пельтье*, состоящие из многих (например, 127) последовательно соединенных термопар, в каждой из которых подключены полупроводники разных типов. На рисунке 42.6 показана схема устройства такой термопары. При пропускании через нее постоянного тока электроны и дырки движутся в одну сторону, охлаждая одну поверхность элемента Пельтье и нагревая другую. Элементы Пельтье широко используются для охлаждения микросхем в электронных устройствах, ПЗС-матриц в приборах ночного видения и цифровых фотоаппаратах, а также в качестве охлаждающих элементов в маломощных автомобильных холодильниках. Элемент Пельтье можно «пустить в обратном направлении», чтобы использовать эффект Зеебека

**Рис. 42.6.** Схема устройства термопары в составе элемента Пельтье: два полупроводника  $n$ - и  $p$ -типа соединены медным проводником и подключены к источнику постоянного тока; стрелками показаны направления тока  $I$  и движения электронов и дырок



и вырабатывать электричество при создании разности температур. В свое время подобными термоэлектрическими генераторами пользовались в отдаленных местах, где не было электричества. Во время Великой Отечественной войны партизаны использовали такое устройство для подзарядки батарей своих раций: разница температур создавалась пламенем костра. Сейчас созданы радиоизотопные термоэлектрические генераторы, в которых тепловая энергия выделяется при естественном распаде радиоактивных изотопов (плутония-238, к примеру) и затем элементами Пельтье преобразуется в электричество. Выходная мощность невысока, КПД мал, но в таких устройствах нет движущихся частей, они не требуют обслуживания и могут использоваться десятилетиями. Такие источники энергии применяются на марсоходе *Curiosity* и в космических зондах, направленных к дальним планетам Солнечной системы.

## Эффект Томсона

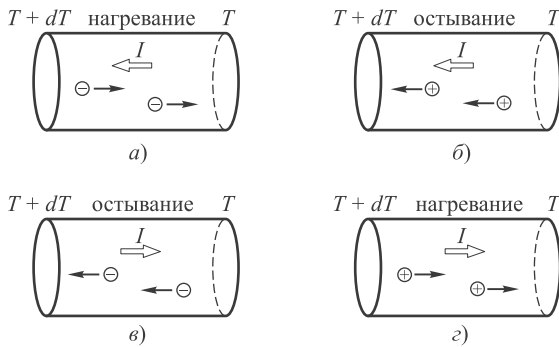
Теплота Пельтье выделяется в контактах на стыке двух разнородных проводников. Но если концы однородного проводника поддерживаются при разных температурах, вдоль него создается градиент температур, так что каждый элемент проводника можно считать отличающимся от соседнего элемента. Поэтому весь проводник можно рассматривать как последовательно соединенные разнородные проводники, отличающиеся друг от друга температурой. На их стыке (т.е. в каждой точке объема проводника) в дополнение к выделению джоулевой теплоты выделяется (или поглощается) **теплота Томсона**, названная так в честь У. Томсона (лорда Кельвина), первооткрывателя и исследователя (1851–1856) этого, уже третьего по счету, термоэлектрического явления. Теплота Томсона  $Q_\tau$ , выделяющаяся на отрезке проводника, пропорциональна протекающему через проводник заряду  $q = It$  и перепаду  $dT$  температур — градиенту температуры вдоль проводника  $\partial T/\partial l$  и на длине отрезка  $dl$ :

$$Q_\tau = -\tau dT It = -\tau \left( \frac{\partial T}{\partial l} dl \right) It. \quad (42.18)$$

Параметр  $\tau$  называется **коэффициентом Томсона**. При положительном коэффициенте Томсона тепло поглощается ( $Q_\tau < 0$ ), если ток течет от холодного конца проводника к горячему, т.е. вдоль градиента температуры ( $IdT/dl > 0$ ). Примером могут служить такие металлы, как медь, серебро, цинк, сурьма, кадмий. При отрицательном коэффициенте Томсона тепло поглощается, если ток течет против градиента температуры, т.е. от горячего конца проводника к холодному. Это происходит в таких металлах, как железо, висмут, кобальт, платина, никель.

Знак постоянной Томсона зависит от относительного влияния двух механизмов, действующих в противоположных направлениях.

Один из них определяется зависимостью средней энергии носителей тока от температуры. Пусть ток направлен вдоль градиента температуры, т.е. он идет в сторону возрастания температуры от холодного конца к горячему. В полупроводниках электронного типа и металлах электроны движутся в обратном направлении, от горячего конца к холодному (рис. 42.7, а). Поскольку средняя энергия электронов на горячем конце выше, они, попадая в более холодную среду, отдают излишек энергии, сталкиваясь с ионами кристаллической решетки и тем самым нагревая материал. В полупроводниках дырочного типа положительно заряженные дырки движутся по направлению тока (рис. 42.7, б) и охлаждают материал. Таким образом, указанный механизм выравнивания

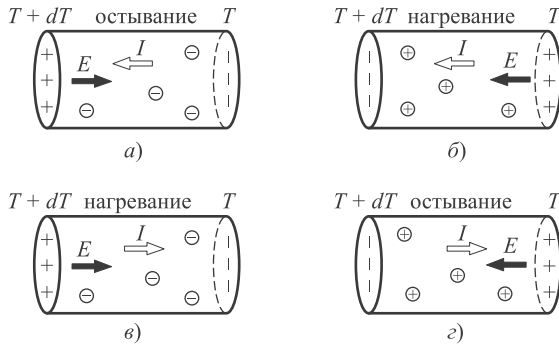


**Рис. 42.7.** Возникновение эффекта Томсона вследствие выравнивания средних энергий носителей тока в металлах и полупроводниках  $n$ -типа ( $a$ ) и полупроводниках  $p$ -типа ( $b$ ); при смене направления тока в полупроводниках обоих типов ( $c$  и  $d$ ) тепловой поток меняет направление — нагревание материала сменяется остыванием и наоборот

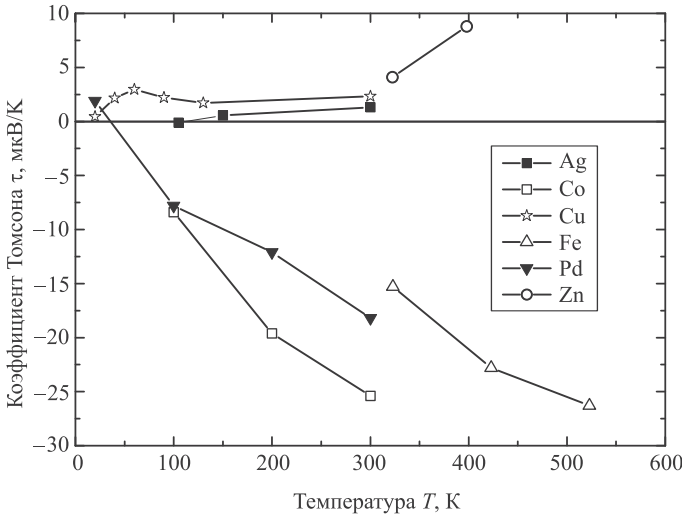
средней энергии носителей тока ведет к отрицательной постоянной Томсона для металлов и полупроводников  $n$ -типа и к положительной для полупроводников  $p$ -типа. При смене направления тока относительно градиента температуры нагревание материала сменяется остыванием, и наоборот (рис. 42.7,  $c$ ,  $d$ ).

Другой механизм связан с диффузией носителей тока и особенно выражен в полупроводниках. В них концентрация носителей тока сильно зависит от температуры, поэтому градиент температуры создает градиент концентрации носителей тока, направленный в ту же сторону. В результате в полупроводниках  $n$ -типа (и металлах) электроны диффундируют от горячего конца проводника к холодному, так что на горячем конце образуется недостаток электронов, т. е. избыток положительного заряда, а на холодном — избыток отрицательного заряда (рис. 42.8,  $a$ ,  $c$ ). Это приводит к появлению электрического поля  $E$ , направленного против градиента температуры, т. е. от горячего конца к холодному. В полупроводниках  $p$ -типа на горячем конце, наоборот, создается избыток отрицательных зарядов, а на холодном — избыток положительных и дополнительное электрическое поле направлено от холодного конца проводника к горячему (рис. 42.8,  $b$ ,  $d$ ). В тех случаях, когда это дополнительное поле направлено противоположно току, оно уменьшает скорость носителей тока (электронов или дырок) и проводник остывает (рис. 42.8,  $a$ ,  $d$ ). Если же направления дополнительного электрического поля и тока совпадают, скорости носителей тока увеличиваются, и проводник нагревается (см. рис. 42.8,  $b$ ,  $c$ ).

Мы действительно убеждаемся, что механизм выравнивания средней энергии носителей тока и диффузионный механизм действуют в противоположные стороны. Какой из двух механизмов возьмет верх и каким окажется в итоге знак коэффициента Томсона  $\tau$ , зависит от вещества проводника и его температуры. Говоря выше о знаке коэффициента Томсона для конкретных материалов, мы имели в виду его значение при комнатной температуре. Зависимость коэффициента Томсона от температуры для различных металлов показана на рис. 42.9. В отличие от коэффициентов Зеебека и Пельтье, определяемых для одного материала относительно другого, коэффициент Томсона можно измерить как абсолютный параметр, характеризующий данное вещество. Зная его, в принципе можно уже рассчитать абсолютные величины других термоэлектрических коэффициентов, поскольку все три термоэлектрических явления связаны между собой.



**Рис. 42.8.** Диффузионный механизм возникновения эффекта Томсона приводит к появлению избыточных зарядов на концах рассматриваемого отрезка проводника и, как следствие, дополнительного электрического поля  $E$ ; если это поле направлено против тока  $I$  (а, г), то оно уменьшает скорости носителей тока, и проводник остывает; если же направления поля и тока совпадают (б, в), то скорости носителей тока возрастают, и проводник нагревается; во всех случаях при смене направления тока тепловой поток тоже меняет направление



**Рис. 42.9.** Зависимость коэффициента Томсона от температуры для разных металлов

Если замкнуть термопару и поддерживать один ее конец при температуре  $T_1 = T + dT$ , а другой при температуре  $T_2 = T$ , то по ней потечет ток. В одном спае за время  $t$  выделится теплота Пельтье  $\Pi_{ab}(T + dT)It$ , в другом поглотится теплота Пельтье  $-\Pi_{ab}(T)It$ . С учетом выделения теплоты Томсона в одном проводнике и ее поглощения в другом, для выделившейся теплоты  $Q_1$  и поглощенной теплоты  $Q_2$  имеем соотношения

$$Q_1 = \Pi_{ab}(T + dT)It + \tau_b dT It, \quad Q_2 = -\Pi_{ab}(T)It - \tau_a dT It. \quad (42.19)$$

Поскольку разность температур  $dT$  мала, выделением джоулева тепла можно пренебречь и считать прохождение термоэлектрического тока обратимым процессом. Тогда к нему можно применить равенство Клаузиуса  $Q_1/T_1 + Q_2/T_2 = 0$  (см. т. 1, разд. 12.7),

выражающее постоянство энтропии в термопаре, состоящей из двух проводников  $a$  и  $b$ :

$$\frac{\Pi_{ab}(T+dT)}{T+dT} - \frac{\Pi_{ab}(T)}{T} + \tau_b \frac{dT}{T} - \tau_a \frac{dT}{T} = 0. \quad (42.20)$$

Здесь  $\Pi_{ab}(T)$  — относительный коэффициент Пельтье для материалов  $a$  и  $b$  при температуре  $T$ . Отсюда получаем соотношение

$$\frac{d}{dT} \left( \frac{\Pi_{ab}(T)}{T} \right) = \frac{\tau_a - \tau_b}{T}. \quad (42.21)$$

Теперь применим первое начало термодинамики (закон сохранения энергии) — полную выделившуюся за время  $t$  теплоту  $\delta Q = Q_1 + Q_2$  приравняем работе термо-ЭДС  $dV = \alpha_{ab} dT$  (см. (42.15)) при прохождении заряда  $q = It$  за это же время:  $dA = qdV = \alpha_{ab} ItdT$ . С учетом (42.19) получаем

$$\alpha_{ab} = \frac{d\Pi_{ab}(T)}{dT} + \tau_b - \tau_a. \quad (42.22)$$

Разность коэффициентов Томсона можно найти из уравнения (42.21):

$$\tau_b - \tau_a = -T \frac{d}{dT} \left( \frac{\Pi_{ab}(T)}{T} \right) = -\frac{d\Pi_{ab}(T)}{dT} + \frac{\Pi_{ab}(T)}{T}. \quad (42.23)$$

Подставляя (42.23) в (42.22), получаем связь относительных коэффициентов Зеебека и Пельтье:

$$\Pi_{ab}(T) = \alpha_{ab} T. \quad (42.24)$$

Полагая относительный коэффициент Пельтье равным разности соответствующих абсолютных коэффициентов ( $\Pi_{ab} = \Pi_a - \Pi_b$ ), мы можем *определить* последние как

$$\frac{d}{dT} \left( \frac{\Pi}{T} \right) = \frac{\tau}{T}, \quad (42.25)$$

где индексы  $a$  и  $b$  уже опущены за ненадобностью, так как соотношение (42.25) для абсолютных коэффициентов Пельтье и Томсона относится уже к одному материалу. Для упрощения обозначений мы опускаем также явное указание на зависимость термоэлектрических коэффициентов от температуры.

Аналогичным образом соотношение (42.24) также можно записать для абсолютных коэффициентов Зеебека и Пельтье:

$$\boxed{\Pi = \alpha T.} \quad (42.26)$$

С учетом (42.26) уравнение (42.25) можно представить в виде связи абсолютных коэффициентов Томсона и Зеебека:

$$\boxed{\tau = T \frac{d\alpha}{dT}.} \quad (42.27)$$

Абсолютный коэффициент Зеебека при данной температуре  $T$  можно найти интегрированием коэффициента Томсона:

$$\alpha = \int_0^T \frac{\tau}{T} dT. \quad (42.28)$$

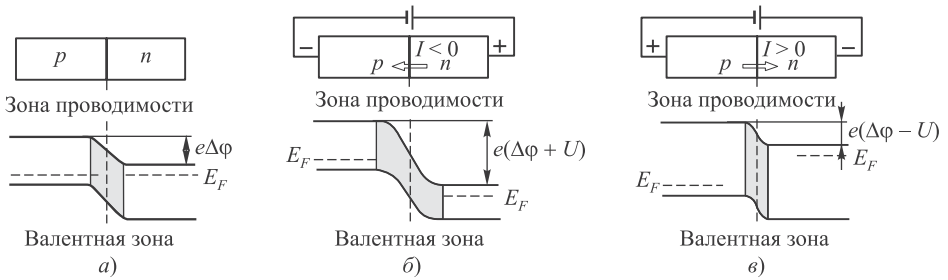
### 42.3 Полупроводниковые выпрямители

При контакте различных полупроводников также выравнивается химический потенциал, а в каждой из сред сохраняется прежнее расположение энергетических зон и уровней относительно химического потенциала (уровня Ферми), за исключением узкого переходного слоя. (По этой же причине, в частности, возникает контактная разность потенциалов при соприкосновении двух металлов.) Сейчас нас интересует переходный слой вблизи контакта двух примесных полупроводников с разным типом проводимости — электронной и дырочной. Напомним, что полупроводник с донорной примесью, в котором основными подвижными носителями заряда являются электроны, называется полупроводником  $n$ -типа. Соответственно полупроводник с акцепторной примесью, где носителями тока являются положительно заряженные частицы (дырки), — это полупроводник  $p$ -типа. Поэтому электронно-дырочный переход обычно называют  $p$ - $n$ -переходом. Электронно-дырочный переход впервые был исследован экспериментально в закиси меди ( $\text{Cu}_2\text{O}$ ) в 1941 г. советским физиком В. Е. Лашкаревым. Поскольку простой механический контакт двух различных полупроводников не может быть достаточно тесным и поэтому не обнаруживает никаких специфических физических свойств,  $p$ - $n$ -переходы создаются внедрением донорных (фосфор, мышьяк и т. п.) и акцепторных (бор, индий и т. п.) примесей в кристалл полупроводника (кремния или германия) с использованием специальных физических приемов. В результате такого *легирования* в кристалле создаются области с разным типом проводимости. На свойствах  $p$ - $n$ -переходов основан принцип действия различных полупроводниковых приборов, поэтому рассмотрим эти свойства подробнее.

#### Контактная разность потенциалов в $p$ - $n$ -переходе

Физический механизм возникновения в  $p$ - $n$ -переходе контактной разности потенциалов основан на существенном различии концентрации электронов и дырок по разные стороны от перехода. Вследствие этого происходит диффузия электронов из  $n$ - в  $p$ -область, а дырок — из  $p$ - в  $n$ -область. Такой встречный процесс диффузии заряженных частиц эквивалентен электрическому току  $I_0$  через  $p$ - $n$ -переход, который течет из  $p$ -области в  $n$ -область. Это направленное движение *основных* носителей заряда называется *диффузионным током*. Встречаясь друг с другом, электроны и дырки интенсивно рекомбинируют, так что пограничный слой будет обеднен подвижными носителями зарядов. Зато в нем останутся некомпенсированными заряды ионов примеси, так что со стороны  $p$ -контакта возникнет слой неподвижных отрицательных ионов акцепторной примеси, а со стороны  $n$ -контакта — положительно заряженный слой ионов донорной примеси. Двойной слой в пограничной области порождает электрическое поле перехода, направленное из  $n$ -области в  $p$ -область. Соответственно, возникает **контактная разность потенциалов**.

Контактная разность потенциалов является причиной встречного направленного движения *неосновных* носителей заряда — дырок из  $n$ -области и электронов из  $p$ -области. Этот «дрейф» заряженных частиц через  $p$ - $n$ -переход направлен



**Рис. 42.10.** Изменение энергетических зон в  $p$ - $n$ -переходе при отсутствии внешнего электрического поля ( $a$ ), при приложении обратного ( $b$ ) и прямого смещения ( $в$ ); направление от  $p$  к  $n$  считается положительным направлением тока  $I$ ; тоном показан запирающий слой

противоположно диффузионному току и называется *дрейфовым током*. Электрическое поле контакта не только является причиной дрейфового тока неосновных носителей, но и препятствует диффузионному току основных носителей заряда. В условиях термодинамического равновесия, которое устанавливается в  $p$ - $n$ -переходе при отсутствии внешней разности потенциалов, диффузионный ток  $I_0$  по величине точно равен дрейфовому току  $-I_0$ , поэтому полный ток через переход равен нулю.

Возникновение контактной разности потенциалов наглядно иллюстрируется зонной диаграммой  $p$ - $n$ -перехода. Рассмотрим сначала рис. 42.10,  $a$ , показывающий изменение расположения энергетических зон в области контакта при отсутствии внешнего электрического поля. В левой части, в области полупроводника  $p$ -типа, уровень Ферми находится у акцепторного уровня, т. е. вблизи потолка валентной зоны, а в области полупроводника  $n$ -типа — вблизи донорного уровня, т. е. у дна зоны проводимости. Соответственно при контакте таких полупроводников дно зоны проводимости и другие зоны слева поднимаются, а справа опускаются так, чтобы уровни Ферми (химические потенциалы) выровнялись. Таким образом создается потенциальный барьер между полупроводниками, т. е. возникает контактная разность потенциалов.

### Сопротивление и односторонняя проводимость $p$ - $n$ -перехода

Область  $p$ - $n$ -перехода обеднена подвижными носителями заряда, так как здесь в результате встречного потока электронов и дырок происходит их интенсивная рекомбинация. По этой причине область  $p$ - $n$ -перехода, несмотря на малую протяженность (порядка  $10^{-6} - 10^{-8}$  м), обладает гораздо большим сопротивлением, чем весь кристалл полупроводника. Эта пограничная область называется *запирающим слоем* (смысл названия станет ясен из дальнейшего рассмотрения).

Если приложить к  $p$ - $n$ -переходу внешнюю разность потенциалов, то ток через переход уже не будет равен нулю из-за смещения уровней Ферми. Когда на  $n$ -контакт подается положительный потенциал, а на  $p$ -контакт — отрицательный (рис. 42.10,  $b$ ), внешнее электрическое поле направлено так же, как и поле,



создаваемое контактной разностью потенциалов (двойным электрическим слоем). Внешнее поле такой полярности называется **обратным смещением**. Оно еще больше препятствует диффузионному току (его называют прямым током), способствует дальнейшему обеднению запирающего слоя основными носителями заряда, оттягивая их из области  $p$ -перехода. В итоге растет и толщина запирающего слоя, и его удельное сопротивление. Оба фактора увеличивают сопротивление запирающего слоя, однако обратный ток (ток неосновных носителей, концентрация которых мала по сравнению с концентрацией основных) при этом несколько возрастает.

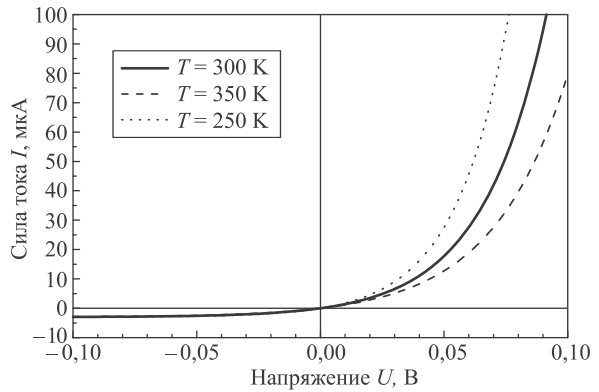
Потенциальный барьер между  $p$ - и  $n$ -областями полупроводника увеличивается за счет относительного смещения уровней Ферми на величину  $-eU$ , где  $U$  — приложенное напряжение, а  $-e$  — заряд электрона (на рис. 42.10, б напряжение отрицательно, поэтому величина  $-eU$  положительна). Неосновные носители заряда — электроны из  $p$ -области, движущиеся слева направо, а также дырки из  $n$ -области, движущиеся в обратном направлении, — свободно скатываются с потенциальной «горки» в  $n$ - и  $p$ -области соответственно. Поэтому через переход продолжает идти небольшой обратный ток  $I_{\leftarrow} = -I_0$ . Движение же основных носителей заряда затруднено, как уже было сказано, из-за обратного направления напряжения смещения. Увеличение потенциального барьера при  $U < 0$  приводит к уменьшению прямого тока в соответствии с Больцмановским фактором:  $I_{\rightarrow} = I_0 \exp(eU/(k_B T))$ . Полный же ток через контакт  $I = I_{\rightarrow} - I_{\leftarrow}$  определяется тогда выражением

$$I = I_0 \left[ \exp\left(\frac{eU}{k_B T}\right) - 1 \right]. \quad (42.29)$$

Из формулы (42.29) видно, что при отсутствии внешнего поля ( $U = 0$ ) ток равен нулю. При увеличении обратного смещения (так что  $e|U|$  превышает несколько величин  $k_B T$ ), экспонента в (42.29) практически равна нулю, отрицательный и небольшой по абсолютной величине ток через переход ( $I \approx -I_0$ ) достигает насыщения, т. е. перестает зависеть от напряжения. При дальнейшем увеличении отрицательного напряжения происходит *электрический пробой  $p$ - $n$ -перехода* — резкий рост тока неосновных носителей. (В области пробоя формула (42.29) уже не работает.) Механизм электрического пробоя состоит в том, что при большом отрицательном напряжении неосновные носители заряда приобретают энергии, достаточные для ионизации атомов в узлах кристаллической решетки полупроводника, что приводит к лавинообразному нарастанию числа электронов. Электрический пробой является обратимым эффектом, он не приводит к повреждению материала полупроводника и исчезает при уменьшении абсолютного значения обратного смещения. Но он может сопровождаться перегревом за счет значительного джоулева тепла, выделяющегося на большом сопротивлении  $p$ - $n$ -контакта, что приводит в таких случаях к необратимым изменениям структуры кристалла (*тепловой пробой*).

При прямом смещении  $U > 0$ , когда к  $p$ -области приложено положительное напряжение, а к  $n$ -области — отрицательное (рис. 42.10, в), внешнее поле направлено против поля контактной разности потенциалов и заставляет основные носители заряда двигаться в область запирающего слоя, тем самым уменьшая

**Рис. 42.11.** Вольт-амперная характеристика  $p$ - $n$ -перехода при разных температурах и постоянном значении  $I_0 = 3$  мкА



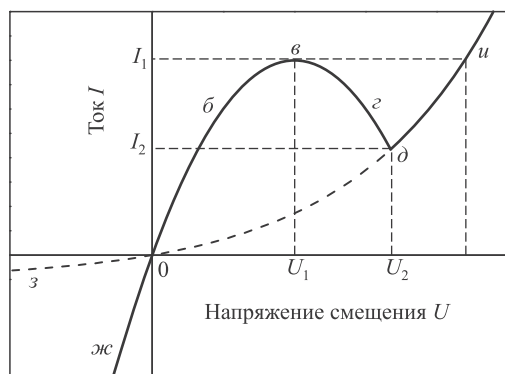
его толщину и удельное сопротивление (а значит, и полное сопротивление слоя). Как следует из формулы (42.29), ток через  $p$ - $n$ -переход возрастает по экспоненциальному закону. Теоретическая вольт-амперная характеристика  $p$ - $n$ -перехода при разных температурах показана на рис. 42.11. При обратном смещении ток практически можно считать равным нулю, поэтому говорят, что  $p$ - $n$ -переход обладает односторонней проводимостью. Благодаря этому свойству  $p$ - $n$ -переход может служить выпрямителем переменного тока и является основным «кирпичиком», использование которого позволило создать различные полупроводниковые устройства — полупроводниковые диоды, транзистор, ячейки солнечных батарей, светодиоды, интегральные схемы. Например, биполярный транзистор состоит из двух таких переходов, соединенных по схеме  $p$ - $n$ - $p$  или  $n$ - $p$ - $n$ . Разные типы такого прибора изобрели (1947–48) американские физики Д. Бардин, У. Браттейн и У. Шокли (Нобелевская премия в 1956 г. за «за исследования полупроводников и открытие транзисторного эффекта») <sup>6</sup>.

## Туннельный диод

**Вырожденными полупроводниками** называются сильно легированные полупроводники, в которых концентрация примесей столь велика ( $10^{24}$ – $10^{26}$  м<sup>-3</sup>), что собственная проводимость материала уже не проявляется, а основную роль играют свойства примесей. Уровень Ферми в вырожденных полупроводниках лежит внутри разрешенных или запрещенных зон, причем его расстояние от границы разрешенной зоны не превышает нескольких величин  $k_B T$ .

В собственном полупроводнике  $n$ -типа валентная зона почти полностью заполнена электронами. Стало быть, зона проводимости почти пуста даже при высоких температурах. В разделе 41.8 мы видели, что при низких температурах в полупроводниках с донорной примесью уровень Ферми находится в запрещенной зоне посередине между донорным уровнем и дном зоны проводимости. При типичных концентрациях примесей порядка  $10^{22}$ – $10^{23}$  м<sup>-3</sup> толщина

<sup>6</sup>Заметим, что Бардин — один из четырех дважды нобелевских лауреатов (кроме него — М. Кюри, биохимик Ф. Сенгер и химик Л. Поллинг) и единственный ученый, получивший две Нобелевские премии по физике.



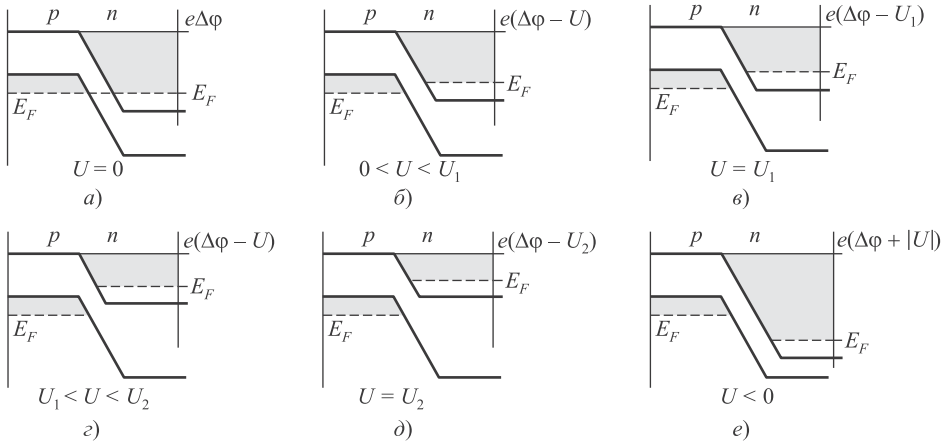
**Рис. 42.12.** Вольт-амперная характеристика туннельного диода; пунктирная линия  $з-0-д$  показывает ход кривой для обычного  $p-n$ -перехода между невырожденными полупроводниками

$p-n$ -перехода в германии составляет около  $1-0,1$  мкм, а электрическое поле невелико:  $10^5-10^6$  В/м.

При увеличении концентрации примеси уровень Ферми приближается к границе зоны проводимости и даже может оказаться внутри нее. Такой полупроводник уже ведет себя как металл (правда, концентрация электронов в нем намного меньше, чем в металле). В полупроводнике  $p$ -типа уровень Ферми также находится внутри запрещенной зоны, но в середине щели между акцепторным уровнем и потолком валентной зоны. При увеличении концентрации примеси уровень Ферми понижается, приближаясь к потолку валентной зоны или даже опускаясь в нее, так что валентная зона становится мало отличающейся от зоны проводимости. Расположение уровней Ферми в разрешенных зонах означает, что дно зоны проводимости  $n$ -области расположено ниже потолка валентной зоны  $p$ -области. Отсюда следует, что для перехода из одной области в другую электронам не нужно сообщать дополнительную энергию: они могут просачиваться на свободные уровни энергии благодаря туннельному эффекту. Толщина переходного слоя уменьшается при увеличении концентрации примесей и в сильно легированных полупроводниках по порядку величины равна  $10-20$  нм, т. е. в 10 или даже в 100 раз меньше, чем толщина переходного слоя при контакте невырожденных полупроводников, что значительно облегчает туннелирование. Соответственно электрическое поле в области контакта вырожденных полупроводников (отношение контактной разности потенциала к толщине слоя) резко возрастает до значений  $10^7-10^8$  В/м.

На основе вырожденных полупроводников японский физик Л. Эсаки создал в 1957 г. новое по тем временам полупроводниковое устройство — **туннельный диод** (Нобелевская премия в 1973 г. совместно с другими учеными, о чем будет сказано в следующей главе)<sup>7</sup>. Вольт-амперная характеристика туннельного диода схематически показана на рис. 42.12. Характерные ее особенности: 1) обратный ток не мал, он имеет тот же порядок, что и прямой ток, и растет с увеличением обратного смещения; 2) при положительном смещении ток сначала растет и

<sup>7</sup>Советские физики Я. И. Френкель и А. И. Иоффе еще в 1932 г., рассматривая свойства выпрямляющего контакта металл-полупроводник, высказали мысль о возможной роли туннельного эффекта. Из-за низкого уровня полупроводниковой техники того времени их гипотеза не могла быть подтверждена в экспериментах.



**Рис. 42.13.** Структура энергетических зон туннельного диода при разных значениях напряжения  $U$ ; пунктирной линией показаны уровни Ферми в  $p$ - и  $n$ -областях; величина  $\Delta\varphi$  — контактная разность потенциалов; заштрихованы участки разрешенных зон, свободные от электронов при нулевой температуре

достигает максимума  $I_1$  при напряжении  $U_1$ , далее идет область ( $z$ ) аномального поведения, где ток падает с ростом напряжения. Эта область отрицательного сопротивления используется для усиления слабых сверхвысокочастотных сигналов. Наконец, при напряжении  $U_2$  ток достигает своего минимального значения  $I_2$ , после чего снова начинает расти, как в обычных  $p$ - $n$ -переходах.

Структура энергетических зон показана на рис. 42.13. При отсутствии напряжения смещения уровни Ферми  $p$ - и  $n$ -областей совпадают (рис. 42.13,  $a$ ), где штриховкой показаны заполненные уровни энергии. Переходы электронов из одной зоны в другую взаимно уравниваются, т. е. токи слева направо и справа налево равны по величине, а полный ток равен нулю.

Если на туннельный диод подано напряжение  $0 < U < U_1$ , то уровень Ферми  $n$ -области повышается, и часть ее заполненной зоны проводимости соответствует по высоте свободным уровням валентной зоны  $p$ -области (рис. 42.13,  $b$ ). Происходит туннелирование электронов из  $n$ -области в  $p$ -область. Туннельный ток, создаваемый этими переходами, намного превышает обычный диффузионный ток. При увеличении прямого напряжения этот туннельный ток растет и достигает максимума в точке  $U = U_1$ , когда уровень Ферми  $n$ -зоны совпадает с потолком валентной зоны  $p$ -области (рис. 42.13,  $в$ ). Максимальный ток  $I_1$  достигается при значениях  $U_1 = 40$ – $50$  мВ для германиевых диодов и при  $U_1 = 100$ – $150$  мВ для диодов на основе арсенида галлия (GaAs) — химического соединения галлия и мышьяка.

При дальнейшем повышении напряжения и относительного смещения уровней Ферми  $n$ - и  $p$ -областей перекрытие заполненных и свободных уровней уменьшается, и туннельный ток падает (рис. 42.13,  $z$ ). Когда дно зоны проводимости  $n$ -области окажется на высоте потолка валентной зоны  $p$ -области, заполненные уровни зоны проводимости будут располагаться «напротив» запрещенной зоны

(рис. 42.13, *д*), и переход электронов при той же энергии окажется невозможным. Туннельный ток должен уменьшиться до нуля, но потенциальный барьер в этом режиме настолько низок, что через *p-n*-переход начинает идти диффузионный ток. При дальнейшем повышении напряжения туннельный диод ведет себя как обычный *p-n*-контакт невырожденных полупроводников.

При подаче на туннельный диод обратного смещения ( $U < 0$ ) уровень Ферми *n*-зоны опускается и свободные уровни зоны проводимости оказываются «напротив» свободных уровней валентной зоны (рис. 42.13, *е*). В этом режиме преобладает поток туннелирующих электронов из валентной зоны *p*-области в зону проводимости *n*-области, т.е. ток имеет обратное направление. Величина обратного тока даже превышает величину прямого тока при напряжениях  $0 < U < U_1$ . Напомним, что обратный ток через *p-n*-переход невырожденных полупроводников очень мал.

Ток в туннельных диодах создается основными носителями заряда и не связан с диффузией через контакт неосновных носителей. Поэтому туннельные диоды обладают очень малой инерционностью. Они отличаются малым потреблением мощности, небольшими габаритами и массой, устойчивостью к радиационному воздействию. Туннельные диоды широко применяются как предварительные усилители, генераторы и высокочастотные переключатели.

## Фотоэлектрические явления в полупроводниках

В разделе 34.2 мы рассмотрели внешний фотоэффект, заключающийся в вырывании электронов из материала проводника электромагнитным излучением. Сходное явление в полупроводниках называется **внутренним фотоэффектом (фотопроводимостью)**. Это явление заключается в том, что при поглощении фотона меняется сопротивление полупроводника вследствие изменения концентрации носителей заряда.

Явление **собственной фотопроводимости** связано с переходом электронов из валентной зоны в зону проводимости. Такое возможно, когда энергии поглощенного фотона хватает на преодоление электроном запрещенной зоны. Если ее ширина равна  $E_g$ , то красная граница собственной фотопроводимости дается соотношением, аналогичным уравнению Эйнштейна (34.8) для внешнего фотоэффекта:

$$\nu_{\text{кр}} = \frac{E_g}{h}. \quad (42.30)$$

Ширина запрещенных зон в широко используемых полупроводниках при  $T = 300$  К равна  $E_g = 0,67$  эВ (германий) и  $E_g = 1,11$  эВ (кремний). Соответствующие длины волн  $\lambda_{\text{кр,Ge}} = 1,85$  мкм и  $\lambda_{\text{кр,Si}} = 1,12$  мкм — это области инфракрасного излучения.

**Примесная фотопроводимость** в случае *n*-полупроводников обусловлена переходом электрона из атома донорной примеси в зону проводимости. Соответственно красная граница примесной фотопроводимости определяется энергией

ионизации атома донорной примеси, т. е. разностью между дном зоны проводимости  $E_c$  и энергетическим уровнем  $E_d$  атома:

$$\nu_{\text{кр}} = \frac{E_c - E_d}{h}. \quad (42.31)$$

В случае полупроводников  $p$ -типа примесная фотопроводимость связана с переходом электрона из валентной зоны в атом акцепторной примеси. На это требуется энергия, превышающая разность между энергетическим уровнем  $E_a$  атома акцепторной примеси и потолком валентной зоны  $E_v$ . Таким образом, красная граница примесной фотопроводимости полупроводника  $p$ -типа записывается как

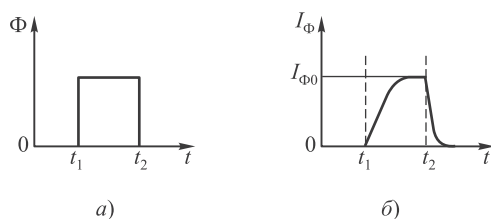
$$\nu_{\text{кр}} = \frac{E_a - E_v}{h}. \quad (42.32)$$

Энергия ионизации атомов донорной примеси и сродство к электрону атомов акцепторной примеси составляют, как уже отмечалось, десятые и даже сотые доли электронвольта. Значит, красная граница примесной фотопроводимости соответствует длинам волн  $\lambda_{\text{кр}} \sim 10\text{--}10$  мкм, т. е. лежит в дальней инфракрасной области электромагнитного спектра.

Появление избыточных *неравновесных носителей заряда*, генерируемых вследствие поглощения полупроводником света, уменьшает сопротивление полупроводника. Концентрация неравновесных носителей зависит от освещенности полупроводника, а не от его температуры, что отличает их от равновесных носителей, концентрация которых определяется именно температурой. Полупроводники, сопротивление которых существенно зависит от освещенности их поверхности, называются *фоторезисторами*. Если подключить фоторезистор к источнику напряжения, через него пойдет слабый *темновой ток*  $I_0$ . При освещении фоторезистора концентрация избыточных носителей заряда будет возрастать, соответственно будет возрастать и ток  $I$  через него. Разность  $I_{\text{Ф}} = I - I_0$  называется *фототоком*. Фототок может превышать темновой ток в сто тысяч или даже в миллион раз.

При достаточно длительном облучении с постоянной интенсивностью устанавливается равновесная для заданной интенсивности концентрация избыточных носителей тока, при которой число появляющихся в единицу времени свободных электронов (дырок) равно числу исчезающих в результате рекомбинации. При слабом освещении фототок пропорционален световому потоку, при увеличении освещенности интенсивность рекомбинации растет быстрее и рост фототока замедляется (зависимость фототока от интенсивности освещения становится нелинейной — фототок насыщается).

Важным отличием фотопроводимости от внешнего фотоэффекта является ее *инерционность*. Дело в том, что требуется некоторое время для установления стационарной концентрации избыточных носителей тока, пока сравниваются скорость фотоионизации и скорость рекомбинации. При внезапном освещении поверхности фоторезистора в нем начинают возникать избыточные носители заряда. Пока их мало, мала и скорость рекомбинации, так что фототок растет. С увеличением концентрации носителей заряда растет и скорость рекомбинации. В какой-то момент наступает равновесие, и фототок перестает расти



**Рис. 42.14.** Интенсивность  $\Phi$  светового импульса, падающего на полупроводник в период времени  $t_1 \leq t \leq t_2$  (а), и соответствующее изменение фототока  $I_\Phi$  (б), происходящее с запаздыванием; величина  $I_{\Phi 0}$  — равновесное значение фототока при данной освещенности (темновой ток не показан, так как он пренебрежимо мал)

и остается постоянным при постоянной интенсивности облучения. Если далее выключить освещение, то при отсутствии вновь порождаемых носителей заряда продолжающаяся рекомбинация (сначала интенсивная, затем ослабевающая) уже имеющихся зарядов ведет к уменьшению тока до нуля. На рисунке (42.14, а) показан график изменения освещенности  $\Phi$  полупроводника при мгновенном включении света в момент времени  $t_1$  и мгновенном же выключении в момент  $t_2$ , а на рис. (42.14, б) — соответствующее этому световому импульсу изменение во времени фототока  $I_\Phi$ . Горизонтальный участок кривой изменения фототока соответствует равновесному значению  $I_{\Phi 0}$  фототока при заданной постоянной освещенности. Инерционность фоторезисторов ограничивает их применение для регистрации меняющихся световых потоков. Тем не менее, они используются для автоматического выключения уличного освещения, в турникетах на входе в метро и многих других областях.

Огромное прикладное значение имеет *вентильный фотоэффект*, или *фотоэффект в запирающем слое*. Полупроводниковые приборы, принцип действия которых основан на этом эффекте, называются *фотоэлементами*. Фоторезистор не имеет области  $p$ - $n$ -перехода, и свойства проводимости для него одинаковы в обе стороны. Но наличие  $p$ - $n$ -перехода меняет все дело. Как уже было сказано, в условиях равновесия диффузионный и дрейфовый токи компенсируют друг друга. При освещении рабочей области фотоэлемента вблизи  $p$ - $n$ -перехода будут возникать неравновесные носители заряда. *Поле перехода разделяет пары неравновесных носителей заряда*. Электроны, возникшие в результате поглощения фотонов в  $n$ -полупроводнике, равно как и дырки, возникшие в  $p$ -полупроводнике, при попытке «пересечь границу» будут возвращены электрическим полем запирающего слоя по месту их «постоянного проживания». Но электроны, возникшие под действием света в  $p$ -полупроводнике, и дырки, появившиеся в  $n$ -полупроводнике, переносятся запирающим слоем в  $n$ - и  $p$ -области соответственно. В результате в  $p$ - $n$ -переходе происходит разделение зарядов, т. е. возникает *фото-ЭДС*, причем  $p$ -область заряжается положительно, а  $n$ -область — отрицательно.

Фотоэлементы непосредственно преобразуют энергию электромагнитного излучения видимого и инфракрасного диапазонов в электрическую энергию. Правда, эффективность фотоэлементов низка — их КПД в лучшем случае составляет

около 20%. Остальная световая энергия идет на нагревание, что понижает эффективность фотоэлементов. И все же в мире достаточно широко используются кремниевые *солнечные батареи*, собранные из отдельных фотоэлементов. Они бывают разного размера — от встроенных в калькуляторы до подвешенных на уличных столбах и размещенных на крышах зданий. Их устанавливают на крышах поездов для освещения вагонов, питания кондиционеров и аварийных систем. Солнечные батареи используются в космосе. Помимо энергетических установок локального применения построены большие солнечные электростанции (СЭС), использующие фотоэлементы. Крупнейшей из них на настоящий момент является «Солнечная звезда» (*Solar Star*), построенная в 2015 г. в поселке Розамонд неподалеку от Лос-Анджелеса (Калифорния, США). Мощность этой СЭС составляет 579 МВт, в ней используется 1,7 млн фотоэлектрических панелей, занимающих площадь около 13 кв. км. «Солнечная звезда» снабжает электричеством 255 тыс. домов. Предыдущие СЭС такого же типа и примерно той же мощности (550 МВт), построенные за год-другой до «Солнечной звезды», использовали около 9 млн панелей, которые занимали площадь 25 кв. км. Прогресс в технологиях налицо. Большие СЭС эксплуатируются также в Греции, Испании и других странах Европы, в Китае и Японии, в Африке и Австралии.

Еще одно фотоэлектрическое явление — возникновение оптического излучения при пропускании электрического тока через *p-n*-переход. Это явление называется *электролюминесценцией*. Создание электромагнитным излучением избыточных носителей заряда в полупроводнике всегда сопровождается обратным процессом рекомбинации. При пропускании тока через *p-n*-переход в прямом направлении — от *p*- к *n*-полупроводнику — электроны из *n*-области направляются к запирающему слою, а через него — в *p*-область. Дырки же, наоборот, уносятся током из *p*-области в *n*-область. В результате в этих областях увеличивается концентрация разноименных зарядов и повышается вероятность их встречи. Это ведет к резкому росту интенсивности рекомбинации. Как мы уже говорили в разд. 41.8, выделяющаяся при рекомбинации энергия может быть унесена разными способами, в том числе и излученным фотоном (*излучательная рекомбинация*). В некоторых полупроводниках рекомбинация имеет, как правило, излучательный характер, в них пропускание прямого тока через *p-n*-переход ведет к оптическому излучению в узком интервале длин волн. На этом принципе работают *светодиоды*. Их излучение не является тепловым, они почти не нагреваются при работе. Явление электролюминесценции было открыто советским изобретателем О. В. Лосевым (1924). Светодиоды используются в бытовой электронике и декоративной светотехнике.

Главу о контактных явлениях уместно завершить хотя бы мимолетным упоминанием *гетеропереходов* — контактов *разнородных* полупроводников (одного или разного типов проводимости) с различными ширинами запрещенных зон. Гетеропереходы находят широкое применение при изготовлении светодиодов, лазеров, фотоэлементов и многих других полупроводниковых устройств. Современные технологии позволяют получать не только двухслойные, но и многослойные гетероструктуры. Примером может служить широко используемое соединение  $\text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$  (здесь  $x$  — доля атомов алюминия, заместивших атомы галлия). В 2002 г. наш соотечественник Ж. И. Алферов и немецкий физик



Г. Крёмер получили Нобелевскую премию «за разработку полупроводниковых гетероструктур, используемых в высокочастотных схемах и оптоэлектронике».

## Контрольные вопросы

1. Объясните причины существования работы выхода.
2. Каково происхождение контактной разности потенциалов?
3. Что такое внешняя и внутренняя контактная разность потенциалов?
4. Объясните возникновение контактной ЭДС в термопаре и ее полярность.
5. Объясните возникновение объемной ЭДС в термопаре.
6. Опишите перестройку энергетических зон при контакте полупроводников  $n$ - и  $p$ -типа. Объясните выпрямляющие свойства такого контакта.
7. Что такое контактная разность потенциалов в  $p$ - $n$ -переходе и чем отличаются диффузионный и дрейфовый токи?
8. Каковы основные отличия характеристик туннельного диода от обычного  $p$ - $n$ -перехода?
9. Опишите перестройку энергетических зон в туннельном диоде в зависимости от напряжения смещения.
10. Почему  $p$ - $n$ -переход называют запирающим слоем?
11. Что такое внутренний фотоэффект?
12. Объясните инерционность внутреннего фотоэффекта.
13. Чем отличается примесная фотопроводимость от собственной фотопроводимости?
14. Что такое фотоэффект в запирающем слое? Объясните принцип действия солнечных батарей. Какая фотопроводимость здесь работает: примесная или собственная?
15. В чем заключается явление электролюминесценции?

## Глава 43

# Макроскопические квантовые явления

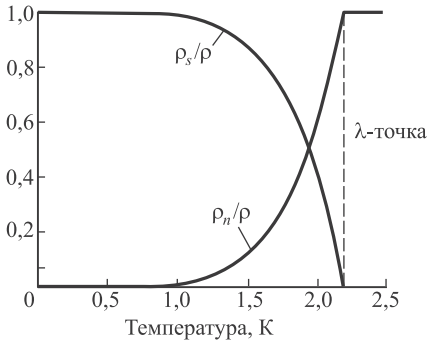
Не следует думать, будто законы квантовой механики важны только при рассмотрении явлений в масштабах атомов и молекул. При низких температурах вблизи абсолютного нуля существуют и макроскопические проявления этих законов. С ними мы и познакомимся в этом разделе.

### 43.1 Сверхтекучесть

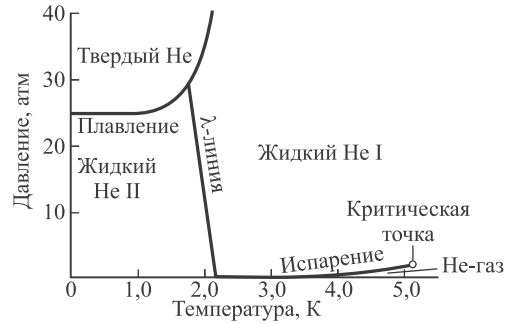
Сверхтекучесть — это свойство жидкого гелия протекать без трения сквозь тонкие капилляры при низких температурах  $T < 2,17$  К. Явление открыто П. Л. Капицей в 1938 г. (Нобелевская премия была присуждена в 1978 г.).

Дальнейшие исследования показали, что в жидком гелии присутствуют два компонента: нормальный и сверхтекучий. Из-за отсутствия вязкости сверхтекучий гелий не оказывает сопротивления погруженным телам. В эксперименте Э. Л. Андроникашвили (1946) стопка металлических дисков, нанизанных на равном расстоянии друг от друга на вертикальную нить, могла совершать крутильные колебания в жидком гелии. Расстояние между дисками было таким малым, что вся жидкость увлекалась вместе с ними и участвовала в колебаниях. Однако при температурах ниже некоторой  $T_\lambda$  (так называемая *лямбда-точка*) частота колебаний резко увеличивалась. Это указывало на то, что не вся жидкость участвует в колебаниях — сверхтекучий компонент дисками не увлекается. Этот эксперимент позволил прямо измерить отношение плотности  $\rho_n$  нормальной компоненты к плотности всей жидкости  $\rho$  и тем самым определить плотность сверхтекучего компонента  $\rho_s$ . Результаты представлены на рис. 43.1. Видно, что при  $T < 1$  К в He II доминирует сверхтекучий компонент.

В сверхтекучем состоянии жидкий гелий проявляет так много необычных свойств, что мы вправе считать сверхтекучий компонент особым состоянием вещества, которое принято обозначать He II (в отличие от обычного He I). Упомянем лишь *эффект фонтанирования*: если опустить в жидкий гелий капилляр и осветить карманным фонарем его нижний конец, то жидкий гелий в этом месте нагреется и доля сверхтекучего компонента упадет. Избыточное давление застав-



**Рис. 43.1.** Отношения плотностей  $\rho_n$  нормального (He I) и  $\rho_s$  сверхтекучего (He II) компонентов к плотности  $\rho$  жидкого гелия в зависимости от температуры ниже  $\lambda$ -точки



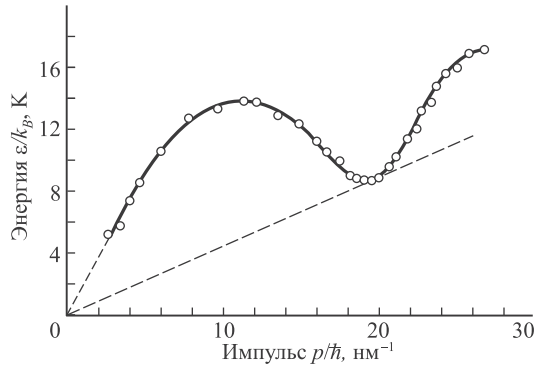
**Рис. 43.2.** Фазовая диаграмма  ${}^4\text{He}$ :  $\lambda$ -линия разделяет нормальное (He I) и сверхтекучее (He II) состояния

ляет втекать в капилляр сверхтекучий компонент, что приводит к появлению фонтанчика высотой до 30–40 см. Такое перетекание гелия от холодного места к горячему необычно: в нормальных жидкостях все происходит как раз наоборот.

Фазовая диаграмма для гелия представлена на рис. 43.2. Фазовый переход He I  $\Rightarrow$  He II происходит на так называемой  $\lambda$ -линии, причем температура перехода уменьшается с ростом давления. Сверхтекучее состояние наблюдалось для изотопа  ${}^4\text{He}$  (сверхтекучесть у другого изотопа  ${}^3\text{He}$  увидели гораздо позже и с большим трудом, но это — другая история). Поскольку протон и нейтрон, входящие в ядро гелия, имеют спин  $1/2$ , то спин ядра  ${}^4\text{He}$  — целый, а  ${}^3\text{He}$  — полуцелый. Здесь возникает разница между бозонами и фермионами. Частицы жидкого  ${}^4\text{He}$  — бозоны, и они могут претерпевать конденсацию Бозе–Эйнштейна. К тому же при типичной концентрации частиц в жидкости  $N/V = 2 \cdot 10^{28} \text{ м}^{-3}$ , массе ядра гелия  $m = 6,7 \cdot 10^{-27} \text{ кг}$  и  $g = 1$  (полный спин ядра равен нулю) формула (40.68) дает для температуры бозе–эйнштейновской конденсации значение  $T_0 = 2,9 \text{ К}$ , что не так далеко от температуры  $\lambda$ -перехода. Представляется весьма соблазнительным связать два необычных состояния вещества — бозе–эйнштейновский конденсат и сверхтекучий компонент. Но все не так просто, как кажется.

## Элементы микроскопической теории сверхтекучести

Дело в том, что жидкость — это не идеальный газ, в ней имеется сильное взаимодействие между частицами. А мы рассматривали конденсат в системе идеального газа бозонов. И надо еще разобраться, почему конденсат должен проявлять сверхтекучесть. Согласно микроскопической теории Н. Н. Боголюбова (1947), нельзя говорить о состоянии отдельных атомов гелия, а следует рассматривать состояние всей системы. Здесь мы впервые встречаемся с *квантовой жидкостью*, с *макроскопическим квантовым эффектом*. При  $T = 0$  система находится в основном состоянии, при повышении температуры возникают тепловые возбуждения, переход на возбужденные уровни энергии. Коллективные возбуждения



**Рис. 43.3.** Зависимость энергии квазичастиц в He II от их импульса (закон дисперсии)

атомов квантуются аналогично колебаниям атомов кристаллической решетки в теории Дебая. Там при квантовании возникали фононы, здесь — квазичастицы, отдельные кванты с энергией  $\varepsilon(\vec{p})$ . Оказалось, что первостепенную роль играет закон зависимости энергии от импульса  $\vec{p}$  — закон дисперсии. В квантовой бозе-жидкости возбуждения (квазичастицы) — это также бозоны с нулевым спином. При низких температурах число квазичастиц мало, и их можно считать невзаимодействующими. В этом случае имеем идеальный газ квазичастиц, который при какой-то критической температуре испытывает конденсацию. Но все, повторим, зависит от закона дисперсии, полный расчет которого очень сложен и до сих пор не выполнен.

Боголюбов вычислил дисперсию слабовозбужденных состояний. В предельных случаях его формула дает

$$\begin{aligned}\varepsilon(\vec{p}) &\sim \frac{p^2}{2m_{\text{эфф}}} \quad (\text{большие импульсы}), \\ \varepsilon(\vec{p}) &\sim up \quad (\text{малые импульсы}),\end{aligned}\quad (43.1)$$

где  $u$  — некая постоянная с размерностью скорости. Энергия в первом случае есть не что иное, как энергия атома гелия (точнее, квазичастицы, которая ведет себя как обычная частица с той разницей, что ее масса может отличаться от массы атома гелия). Энергия же во втором предельном случае аналогична энергии акустического кванта (фонона), где  $u$  — скорость звука. Ход кривой спектра возбуждений в He II, полученного в экспериментах по рассеянию нейтронов, показан на рис. 43.3 из книги<sup>1</sup>. Обработка экспериментальных данных привела к следующим численным результатам. Дисперсия при малых импульсах описывается линейной функцией  $\varepsilon(p) = up$ , где скорость звука  $u = 239$  м/с. Вблизи локального минимума кривая дисперсии описывается функцией

$$\varepsilon(p) = \Delta + \frac{(p - p_0)^2}{2m_{\text{эфф}}}, \quad (43.2)$$

где  $p_0/\hbar = 19,2$  нм<sup>-1</sup>,  $\Delta/k_B = 8,65$  и эффективная масса квазичастицы  $m_{\text{эфф}} = 0,16m_4$ , где  $m_4$  — масса атома <sup>4</sup>He. Отсюда следует уравнение прямой, касательной к кривой дисперсии:  $\varepsilon' = u^*p$ ,  $u^* \approx 14$  м/с (показана штриховой линией на рис. 43.3). Значение параметра  $u^*$  мы обсудим ниже.

Таким образом, при  $T \neq 0$  жидкий гелий состоит из конденсата квазичастиц и идеального газа квазичастиц с энергией  $\varepsilon(\vec{p})$ . Атомы гелия нельзя разделить

<sup>1</sup>Д. Р. Тилли, Дж. Тилли. Сверхтекучесть и сверхпроводимость. М.: Мир, 1977. С. 51.

на «образующие конденсат» и «порождающие свободные квазичастицы»: все они принимают участие в образовании как основного конденсатного состояния, так и идеального газа. Газ квазичастиц может рассматриваться как нормальный компонент жидкого гелия, а конденсат — как сверхтекучий. Подобная феноменологическая теория двухкомпонентной жидкости была развита Л. Д. Ландау (1941). В 1962 г. Ландау была присуждена Нобелевская премия «за новаторские теории конденсированных сред, в особенности жидкого гелия».

## Сверхтекучесть конденсата

Мы разобрались, почему в жидкости может появиться идеальный газ: свободные частицы газа — это не сами атомы гелия, а возбуждения основного состояния (конденсата). Теперь надо понять происхождение сверхтекучести конденсата.

Сначала напомним результаты классической механики, касающиеся перехода к движущимся системам отсчета. Пусть дана система отсчета  $K$ , в которой радиусы-векторы частиц с массами  $m_i$  обозначены  $\vec{r}_i$ . Импульс такой системы есть  $\vec{P} = \sum_i m_i \vec{v}_i$ , а кинетическая энергия  $E = \sum_i m_i v_i^2 / 2$ . Пусть дана также система отсчета  $K'$ , положение начала координат которой задается вектором  $\vec{R}$ ; скорость  $K'$  относительно  $K$  есть  $\vec{V} = \dot{\vec{R}}$ . В движущейся системе координат радиусы-векторы частиц есть  $\vec{r}'_i = \vec{r}_i - \vec{R}$ . Отсюда немедленно следуют выражения для полного импульса и энергии частиц в движущейся системе координат ( $M$  — полная масса всех частиц)

$$\vec{P}' = \vec{P} - M\vec{V}, \quad E' = E - \vec{V} \cdot \vec{P} + \frac{MV^2}{2}. \quad (43.3)$$

Это нерелятивистские преобразования Галилея, известные из классической механики. Следуя рассуждениям Ландау, применим теперь эти формулы к невозбужденному конденсату, текущему по трубке со скоростью  $\vec{V}$ . В лабораторной системе отсчета  $K$  энергия и импульс конденсата выражаются очевидными формулами

$$\vec{P}_1 = M\vec{V}, \quad E_1 = E_0 + \frac{MV^2}{2}, \quad (43.4)$$

где  $E_0$  — энергия покоящегося конденсата в основном состоянии. Возьмем систему  $K'$ , движущуюся со скоростью конденсата  $\vec{V}$ , так что в ней конденсат покоится. Преобразования Галилея (43.3) тогда дают

$$\vec{P}'_1 = 0, \quad E'_1 = E_0. \quad (43.5)$$

Эти результаты очевидны, и мы напомним только, что все это относится к конденсату в основном состоянии. В системе  $K'$  он покоится, а мимо него со скоростью  $-\vec{V}$  движется трубка. Если есть вязкость, то она проявится в диссипации энергии. Это не может сразу произойти во всей жидкости: сначала возбудятся отдельные внутренние движения, появятся квазичастицы типа фононов. Пусть возникло возбуждение с импульсом  $\vec{p}$  и энергией  $\varepsilon(\vec{p})$ . Для возбужденного конденсата в  $K'$  имеем вполне очевидные соотношения

$$\vec{P}'_2 = \vec{p}, \quad E'_2 = E_0 + \varepsilon(\vec{p}). \quad (43.6)$$

Как эти формулы будут выглядеть в лабораторной системе отсчета? Чтобы перейти назад в  $K$ , используем обратные преобразования Галилея

$$\vec{P}_2 = \vec{P}'_2 + M\vec{V}, \quad E_2 = E'_2 + \vec{V} \cdot \vec{P}_2 - \frac{MV^2}{2}, \quad (43.7)$$

откуда находим:

$$\vec{P}_2 = M\vec{V} + \vec{p}, \quad E_2 = E_0 + \frac{MV^2}{2} + E(\vec{p}) + \vec{V} \cdot \vec{p}. \quad (43.8)$$

Сравнивая формулы (43.4) и (43.8), находим изменение энергии движущегося конденсата в лабораторной системе  $K$  при переходе из основного состояния в возбужденное:

$$\Delta E = E_2 - E_1 = \varepsilon(\vec{p}) + \vec{V} \cdot \vec{p}. \quad (43.9)$$

Образование квазичастицы энергетически выгодно, если  $\Delta E < 0$ . В самом благоприятном для этого случае импульс  $\vec{p}$  антипараллелен скорости  $\vec{V}$ , и изменение энергии конденсата равно  $\Delta E = \varepsilon(\vec{p}) - pV$ . Отсюда следует, что для торможения жидкости необходимо выполнение условия

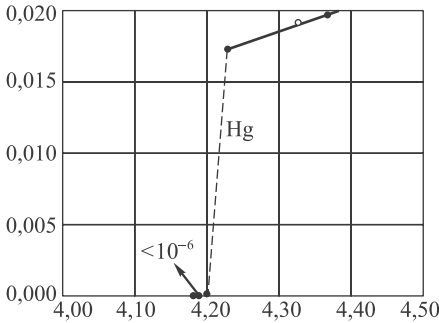
$$V > \frac{\varepsilon(\vec{p})}{p} \geq \min \frac{\varepsilon(\vec{p})}{p} = u^*. \quad (43.10)$$

Здесь мы ввели обозначение  $u^*$  для параметра с размерностью скорости — отношения минимума энергии возбуждения к импульсу. Этот параметр пропорционален тангенсу угла наклона штриховой линии на рис. 43.3. Если спектр возбуждений таков, что  $u^* > 0$ , как это и показано, то условие торможения не выполняется при слабых возбуждениях: как говорят, в спектре возбуждения имеется *энергетическая щель* — некое расстояние между энергией возбуждения и энергией основного состояния. Это значит, что при малых скоростях  $V < u^*$  трение невозможно. Иными словами, мы получили явление сверхтекучести (на рис. 43.3  $u^* \approx 14$  м/с). Смысл полученного результата в том, что конденсат — это коллективный эффект, реагирующий на выпадение одного из своих членов повышением энергии, в то время как трение должно энергию понижать. Закон дисперсии, изображенный на рис. 43.3, таков, что условие сверхтекучести выполнено. Для обычных бозе-частиц с  $E = p^2/(2m)$  имеем  $\min(\varepsilon/p) = 0$  и условие сверхтекучести не выполнено. Вывод: обычные частицы даже в конденсатном состоянии сверхтекучестью не обладают, конденсат не образует связанного коллектива и не реагирует на замедление отдельных частиц.

Еще раз подчеркнем: HeII — это не смесь различных веществ, и его компоненты невозможно выделить по отдельности. Нельзя представлять это так, будто одни атомы гелия принадлежат нормальному компоненту жидкого гелия, а другие — сверхпроводящему компоненту. Все атомы гелия одинаковы, и было бы правильнее говорить, что ниже лямбда-точки ( $T < T_\lambda$ ) в жидком HeII одновременно сосуществуют два вида движения — нормальное и сверхтекучее, причем каждое из них характеризуется своей скоростью, плотностью, вязкостью.

## 43.2 Сверхпроводимость

В 1911 г. голландский физик Г. Камерлинг-Оннес открыл удивительное явление. При температуре около 4 К электрическое сопротивление ртути уменьшалось до нуля. Уменьшение сопротивления металла по мере приближения температуры к абсолютному нулю ожидалось согласно электронной теории металлов, но здесь обнаружилось неожиданное — сопротивление уменьшалось не непрерывно, а скачком. В бумагах Камерлинг-Оннеса нашлась запись: «При 4 К рост сопротивления не замечен. При 4,05 К также ничего. При 4,12 К сопротивление начало появляться». В дальнейшем он установил, что переход в сверхпроводящее состояние очень узок: повышение температуры всего на 0,1 К выше критической увеличивает сопротивление ртути в 400 раз. На рисунке 43.4 показан график зависимости сопротивления ртути от температуры из оригинальной работы Камерлинг-Оннеса. Из графика видно, что при понижении температуры вплоть до значения 4,2 К сопротивление уменьшается по линейному закону, как это и



**Рис. 43.4.** На графике зависимости сопротивления (в омах) от температуры (в кельвинах) из оригинальной работы Камерлинг-Оннеса виден переход ртути в сверхпроводящее состояние при температуре  $T < 4,20$  К

следует из квантовой теории электропроводности металлов, однако потом резко падает почти до нулевого значения.

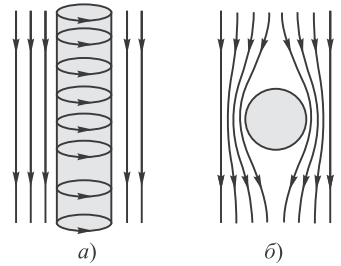
Впоследствии явление *сверхпроводимости* было обнаружено им у олова, свинца, таллия и других веществ. В 1913 г. Камерлинг-Оннес был удостоен Нобелевской премии «за исследования свойств вещества при низких температурах, что привело, в числе прочего, к получению жидкого гелия». В настоящее время сверхпроводящее состояние обнаружено у многих металлов и сплавов.

## Эффект Мейснера и критические значения температуры и магнитного поля

Свойства сверхпроводящего состояния вещества были выявлены в последующих многочисленных экспериментах. Важнейшим из этих свойств (и первым по хронологии открытия) является наличие критической температуры  $T_{c0}$ , при которой возникает сверхпроводящее состояние. Для большинства чистых металлов, являющихся хорошими проводниками в обычных условиях, критические температуры составляют несколько кельвинов.

Около двух десятков лет после открытия сверхпроводники считались идеальными проводниками, т. е. обычными проводниками, какие мы рассматривали в классической теории электричества, но сопротивление которых по какой-то причине становится равным нулю при  $T < T_{c0}$ . Представим себе бесконечный цилиндр, сделанный из такого (гипотетического) идеального «классического» проводника (на самом деле в данном контексте форма проводника особой роли не играет). Охладим его до температуры ниже критической  $T_{c0}$  для данного материала и после этого включим магнитное поле  $H$ , направленное вдоль оси цилиндра (рис. 43.5, а). По поверхности проводника в силу закона Ленца потекут токи, создающие свое магнитное поле, направленное противоположно внешнему, так что магнитное поле внутри проводника будет равно нулю. Это следует из уравнений Максвелла. Действительно, при наложении внешнего поля изменение магнитной индукции внутри образца порождает вихревое электрическое поле (см. (32.18)):

$$\operatorname{rot} \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \quad (43.11)$$



**Рис. 43.5.** Цилиндрический сверхпроводник в магнитном поле, направленном вдоль его оси (а) и поперек нее (б)

Но из уравнения связи (32.19)  $\vec{E} = \vec{j}/\sigma$  следует, что при бесконечно большой удельной проводимости среды (нулевом удельном сопротивлении  $\rho = 1/\sigma$ ) должно быть равно нулю и электрическое поле. Следовательно, производная в правой части (43.11) равна нулю, т. е. магнитная индукция  $\vec{B}$  остается постоянной в идеальном проводнике. Поскольку магнитное поле внутри проводника при отсутствии внешнего поля  $\vec{H}$  было равно нулю, оно останется равным нулю и после его включения. Как следствие, вблизи идеального проводника внешнее поле параллельно его поверхности. Это вытекает из непрерывности нормальных составляющих вектора магнитной индукции на границе раздела двух сред (см. формулу (21.21) в т. 1). Коль скоро магнитная индукция в идеальном проводнике равна нулю, из уравнения Максвелла  $\text{rot } \vec{B} = \mu_0 \vec{j}$  (теорема о циркуляции магнитного поля при отсутствии тока смещения) следует, что и ток в толще идеального проводника должен быть равен нулю, т. е. все токи текут по его поверхности. (Следует оговорить, что токи текут все же не буквально по поверхности идеального проводника, а в тонком слое вблизи нее.)

Но представление о сверхпроводнике как об идеальном проводнике с нулевым сопротивлением рухнуло в 1933 г. после открытия немецкими физиками В. Мейснером и Р. Оксенфельдом эффекта, заключающегося в выталкивании слабого магнитного поля из толщи сверхпроводника (это явление носит название *эффекта Мейснера*). Однако мы только что показали, что и в идеальный «классический» проводник магнитное поле тоже не проникает, так в чем же проблема? А проблема в том, что мы сначала охладили проводник ниже критической температуры, а потом наложили магнитное поле, которое не смогло проникнуть в его толщу. Если же поступить наоборот и поместить «теплый» идеальный проводник во внешнее магнитное поле, то оно спокойно проникнет в его толщу. А при охлаждении «классического» проводника до сверхпроводящего состояния поле так и останется «замороженным» внутри него согласно тому же самому правилу Ленца: при любом изменении магнитного поля в толще идеального проводника возникнут незатухающие токи, этому препятствующие. Таким образом, при данной температуре  $T < T_c$  и данном значении внешнего магнитного поля идеальный «классический» проводник будет находиться в разном состоянии в зависимости от пути, по которому его в это состояние привели. Если внешнего магнитного поля сквозь идеальный проводник изначально не было, то при приближении магнита поле внутри него так и останется нулевым. Если магнит лежал на идеальном проводнике с нулевым сопротивлением, создавая



сквозь него какой-то магнитный поток, то при удалении магнита поле в проводнике также не изменится и не будет равно нулю. А суть эффекта Мейснера как раз и заключается в том, что *слабое магнитное поле не проникает в сверхпроводник при любом пути перехода образца в сверхпроводящее состояние в присутствии внешнего магнитного поля.*

За счет эффекта Мейснера внутренние области сверхпроводника полностью «экранированы» от внешнего магнитного поля. Это означает, что вектор магнитной индукции  $\vec{B}'$  собственного поля поверхностных токов сверхпроводника равен по величине и противоположен по знаку вектору магнитной индукции  $\vec{B}_0$  внешнего поля, так что их сумма равна нулю. Если вспомнить формулу  $\vec{B}' = \chi_m \vec{B}_0$  (см. т. 1, уравнение (21.2)), то можно сказать, что магнитная восприимчивость сверхпроводника  $\chi_m = -1$ , т.е. *сверхпроводник представляет собой не только идеальный проводник с нулевым сопротивлением, но и идеальный диамагнетик.*

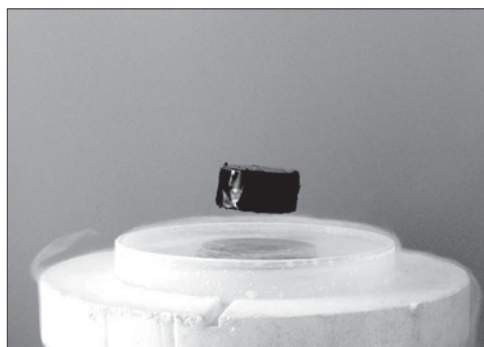
Заметим также, что теорема Гаусса в применении к магнитному полю приводит к выводу о непрерывности нормальной составляющей вектора магнитной индукции на границе раздела сред (см. т. 1, разд. 21.5). Это вытекает из уравнения Максвелла, утверждающего отсутствие в природе магнитных зарядов, и никак не связано с поверхностными токами. Так как магнитное поле внутри сверхпроводника всегда равно нулю, равна нулю и нормальная составляющая магнитного поля вне его. Иными словами, магнитное поле вблизи сверхпроводника всегда параллельно его поверхности (рис. 43.5, б).

Эффект Мейснера демонстрируется опытом по левитации магнита над сверхпроводником (или сверхпроводника над постоянным магнитом). При смещении вверх или вниз магнит возбуждает в сверхпроводнике индукционные токи, которые порождают магнитное поле, препятствующее перемещению магнита. Таким образом, магнит парит над сверхпроводящей подложкой (рис. 43.6). Будучи идеальным диамагнетиком, сверхпроводник выталкивается магнитным полем, что и приводит к левитации магнита над сверхпроводником или сверхпроводника над постоянным магнитом<sup>2</sup>. На этом принципе функционирует магнитоплан, или маглев (от англ. *magnetic levitation*) — поезд на магнитной подушке, способный развивать скорость до 600 км/ч (рис. 43.7). О различных вариантах магнитоплана в разных странах можно прочитать в Интернете (<http://tinyurl.com/ptw6rbf>).

Разрушение сверхпроводящего состояния происходит не только при повышении температуры выше критической, но и в сильном магнитном поле, напряженность которого превышает некоторое критическое значение  $H_{c0}$ . Это критическое значение зависит от температуры сверхпроводника и с хорошей точностью описывается эмпирической параболической зависимостью

$$H_c = H_{c0} \left[ 1 - \left( \frac{T}{T_{c0}} \right)^2 \right]. \quad (43.12)$$

<sup>2</sup>Иногда опыт с парящим магнитом называют «гробом Магомета». Название связано с бытовавшей в средневековой Европе легендой, будто гроб пророка Магомета (Мухаммеда) висит в воздухе без всякой опоры. Христианские писатели сами летающий гроб не видели, но объясняли его «левитацию» тем, что гроб сделан из железа и искусно размещен между двумя магнитами. В исламских источниках летающий гроб пророка не упоминается.



**Рис. 43.6.** Магнит, висящий над высоко-температурным сверхпроводником, охлажденным жидким азотом до  $-196^\circ\text{C}$  (фото из Википедии)



**Рис. 43.7.** Поезд на магнитной подушке (маглев), связывающий г. Шанхай (Китай) с аэропортом. Дорога протяженностью 30 км была открыта в 2002 г., поезд развивает скорость до 450 км/ч (фото из Википедии)

Здесь  $H_{c0}$  — критическое значение напряженности магнитного поля при нулевой температуре. Для каждого значения температуры  $T < T_{c0}$  существует свое значение критического магнитного поля. И наоборот, для каждого значения магнитного поля  $H < H_{c0}$  существует своя критическая температура перехода в сверхпроводящее состояние:

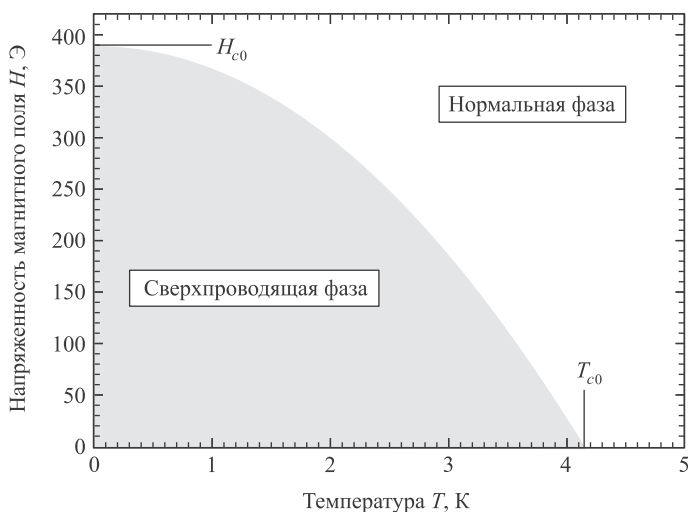
$$T_c = T_{c0} \sqrt{1 - \frac{H}{H_{c0}}}. \quad (43.13)$$

Очевидно, что существование критического поля влечет за собой значение критического тока, который может пройти через сверхпроводник. Большие токи создадут магнитное поле с индукцией выше критической, и сверхпроводящее состояние будет разрушено.

На рисунке 43.8 показан график зависимости критического значения магнитного поля от температуры для ртути.

## Сверхпроводники 2-го рода и вихри Абрикосова

Все сказанное выше относится к так называемым *сверхпроводникам 1-го рода*. Однако существуют *сверхпроводники 2-го рода*, в которых для данной температуры имеются два критических значения напряженности магнитного поля — верхнее  $H_{c1}$  и нижнее  $H_{c2}$ . При полях  $H > H_{c1}$  сверхпроводник находится в нормальной (не сверхпроводящей) фазе, в слабых полях  $H < H_{c2}$  он переходит в сверхпроводящую фазу, а при промежуточных значениях поля  $H_{c2} < H < H_{c1}$  возникает *смешанное состояние*, в котором сверхпроводник пронизан нитеобразными вихрями (их называют *вихрями Абрикосова*, рис. 43.9). Вихри Абрикосова расположены параллельно приложенному магнитному полю и образуют плотно упакованную треугольную решетку. Каждый такой

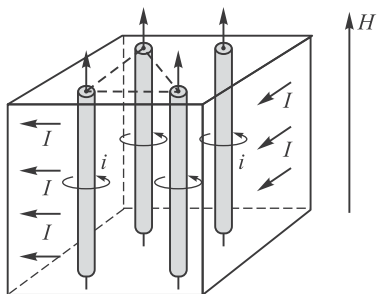


**Рис. 43.8.** Фазовая диаграмма для сверхпроводника 1-го рода (ртути):  $T_{c0} = 4,153$  К,  $H_{c0} = 390$  Э. Значения напряженности магнитного поля показаны в эрстедах; связь с единицей напряженности в СИ дается соотношением  $1 \text{ Э} = 79,5775 \text{ А/м}$ . В вакууме напряженности поля  $H = 1 \text{ Э}$  соответствует магнитная индукция  $B = 10^{-4} \text{ Тл}$

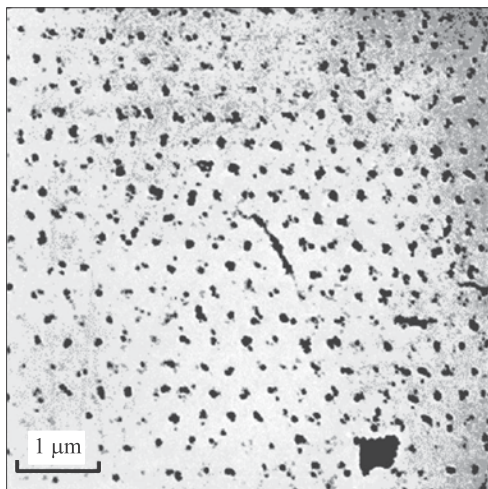
вихрь состоит из центральной сердцевинки малого (до 100 нм) радиуса, окруженной круговыми незатухающими токами. В сердцевинах материал находится в нормальном состоянии, а их радиусы зависят от средней длины свободного пробега электронов, которая в свою очередь зависит от температуры образца. При приближении температуры к критическому значению радиусы сердцевин стремятся к бесконечности, они захватывают весь материал и сверхпроводимость исчезает. Внутри сердцевин имеется магнитное поле, отсутствующее в толще сверхпроводника. Оно образуется круговыми незатухающими токами, текущими по внутренним поверхностям сверхпроводника вокруг сердцевин. Это магнитное поле параллельно внешнему магнитному полю. Стало быть, направления токов в вихрях противоположно направлению токов, текущих по поверхности всего образца и обеспечивающих отсутствие магнитного поле в его толще вне вихрей. Резкой границы между сердцевинкой с нормальным веществом и остальным сверхпроводником не существует — переход размыт, токи и здесь текут в тонком слое на границе между нормальным и сверхпроводящим материалами. Впервые вихри Абрикосова наблюдались экспериментально в 1967 г. (рис. 43.10), через 10 лет после их теоретического предсказания А. А. Абрикосовым в 1957 г.

Фазовая диаграмма для сверхпроводника 2-го рода (конкретно ниобия) показана на рис. 43.11. Существование сверхпроводников 2-го рода было теоретически предсказано в СССР А. А. Абрикосовым (Нобелевская премия 2003 г. совместно с В. Л. Гинзбургом и британско-американским ученым Э. Легеттом «за новаторский вклад в теорию сверхтекучести и сверхпроводимости»).

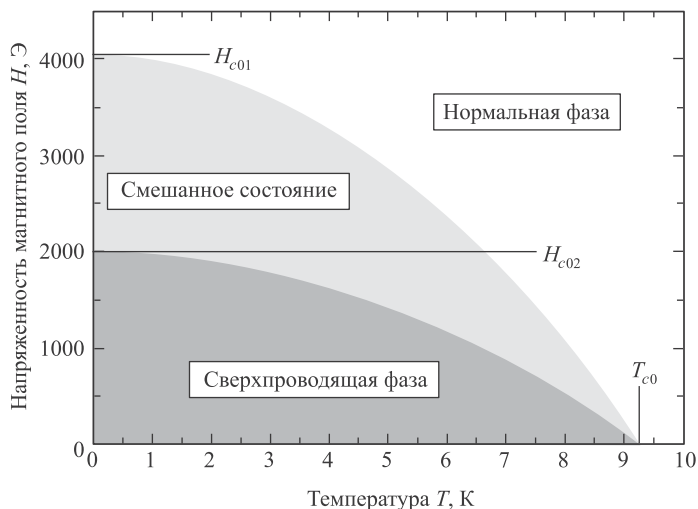
В 1986 г. в исследовательской лаборатории ИВМ в Цюрихе (Швейцария) был открыт керамический материал (барий–лантан–медный оксид), переходящий в



**Рис. 43.9.** По поверхности сверхпроводника, помещенного в магнитное поле  $H$ , текут поверхностные токи  $I$ , компенсирующие магнитное поле в толще материала. В сверхпроводниках второго рода (*внутри*) имеются сердцевинки из нормально-го материала, которые пронизывает магнитный поток, создаваемый вихревыми токами  $i$ , текущими вокруг сердцевин в направлениях, обратных токам  $I$



**Рис. 43.10.** Первое наблюдение вихрей Абрикосова: правильная треугольная решетка видна на поверхности сверхпроводника 2-го рода (свинец—индий при  $T = 1,1$  К) из работы *U. Essmann and H. Trauble. Phys. Lett., 24A, 526, 1967*



**Рис. 43.11.** Фазовая диаграмма для сверхпроводника 2-го рода (ниобия):  $T_{c0} = 9,25$  К, критические значения полей при  $T = 0$  равны  $H_{c01} = 4040$  Э и  $H_{c02} = 1980$  Э

сверхпроводящее состояние при температуре  $T_{c0} = 35 \text{ K}$ , что на 12 градусов выше, чем самые высокие температуры перехода известных ранее сверхпроводников. Авторы открытия — К. Мюллер и Г. Беднорц — получили Нобелевскую премию «за важный прорыв, выразившийся в открытии сверхпроводимости керамических материалов». Последовавшие бурные исследования явления, названного **высокотемпературной сверхпроводимостью** (ВТСП), привели к открытию многих соединений, переходящих в сверхпроводящую фазу при «азотных» температурах. Температура кипения азота составляет 77 К или  $-196^\circ\text{C}$ , а рекорд принадлежит сейчас соединению с температурой перехода  $T_{c0} = 135 \text{ K}$ ; под давлением критическая температура повышается до 165 К ( $-109^\circ\text{C}$ ). Производство жидкого азота, составляющего основную долю земной атмосферы, гораздо дешевле, чем производство жидкого водорода (температура кипения 20 К) и тем более жидкого гелия (температура кипения 4,2 К), с помощью которых получали сверхпроводящее состояние раньше. Поэтому открытие керамических сверхпроводников вселило большие надежды на многочисленные технические приложения ВТСП и даже мечты о материалах, теряющих сопротивление при комнатных температурах. К сожалению, керамики оказались малотехнологичными, и, например, мощные сверхпроводящие магниты Большого адронного коллайдера охлаждаются жидким гелием до температуры 1,9 К. Наибольшие технологические применения имеют сплавы ниобия, сверхпроводящие состояния которых выдерживают высокие значения магнитных полей.

Значения критических параметров для ряда веществ показаны в табл. 43.1.

Для чистого сверхпроводника удельное сопротивление в сверхпроводящем состоянии равно нулю только для постоянного тока. При прохождении переменного тока сверхпроводник оказывает сопротивление, в нем выделяется джоулево тепло, как и в обычном проводнике. Это наблюдение привело к появлению первой теории сверхпроводимости — так называемой двухжидкостной модели сверхпроводника голландских физиков Х. Казимира и К. Я. Гортера (1934), согласно которой в сверхпроводнике существует две группы носителей заряда (электронов): сверхпроводящая и нормальная. Носители заряда сверхпроводящей группы не испытывают сопротивления кристаллической решетки. Ток, создаваемый носителями заряда второй группы, подчиняется обычным законам электромагнетизма: Ома, Джоуля—Ленца и др. Представим себе, что к концам сверхпроводника приложено напряжение, создающее в нем электрическое поле. Оно ускоряет электроны обеих групп, но как только ток примет равновесное постоянное значение, электрическое поле внутри проводника исчезнет (в противном случае скорость сверхпроводящих электронов возрастала бы до скорости света, а вместе с ней рос бы и ток, что физически бессмысленно). При отсутствии электрического поля нормальные электроны не движутся и тока не создают. Ток создается лишь движением по инерции сверхпроводящих электронов, не испытывающих никакого сопротивления. Если же к сверхпроводнику приложено переменное напряжение, то электрическое поле ускоряет носители заряда обеих групп, и у сверхпроводника появляется некоторое сопротивление. Легко оценить частоту  $\nu$  переменного поля, при котором проявится этот эффект: энергия кванта электромагнитного поля должна быть по порядку величины равна тепловой энергии электронов, т. е.  $h\nu = k_B T_{c0}$ . Принимая  $T_{c0} \approx 1 \text{ K}$ , получаем оценку

**Таблица 43.1.** Критические значения температуры и напряженности магнитного поля для некоторых металлов и сплавов; в левой половине таблицы показаны сверхпроводники 1-го рода, в правой — сверхпроводники 2-го рода (среди них материал германит ниобия ( $\text{Nb}_3\text{Ge}$ ), бывший до 1986 г. рекордсменом среди сверхпроводников)

Вещество	$T_{c0}$ , К	$H_{c0}$ , Э	Вещество	$T_{c0}$ , К	$H_{c2}$ , Э
Al	1,180	104,9	Tc	7,8	1410
Hg	4,153	390	V	5,46	1167
In	3,415	289	Nb	9,3	1980
Pb	7,200	803,4	NbTi	9,2	150 000
Sn	3,722	308	$\text{Nb}_3\text{Sn}$	18,0	245 000
Ta	4,46	831	$\text{Nb}_3\text{Ge}$	23,2	370 000
$\text{V}_3\text{Ga}$	14,5	350 000	$\text{Nb}_3\text{AlGe}$	20,7	440 000
$\text{Nb}_3\text{Sn}$	18,0	250 000	$\text{PbMo}_6\text{S}_6$	15	600 000

$\nu \approx 1,38 \cdot 10^{-23}/6,63 \cdot 10^{-34} \approx 2 \cdot 10^{10}$  Гц, что соответствует сантиметровому (СВЧ) диапазону электромагнитных волн.

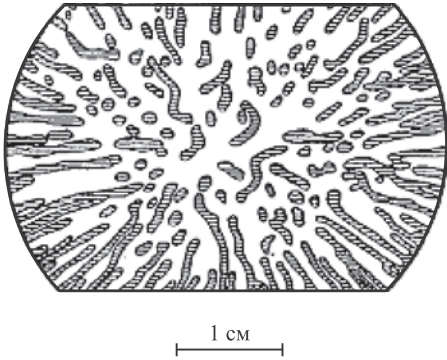
### Промежуточное состояние

Вернемся к рисунку 43.5, *a*, который изображает теперь цилиндрический сверхпроводник 1-го рода, охлажденный до температуры  $T$  ниже критической и помещенный в магнитное поле  $H$ , параллельное его оси. При увеличении поля до значений  $H > H_c$  сверхпроводящее состояние разрушается и цилиндр превращается в нормальный металл.

Но проводники произвольной формы искажают внешнее поле вблизи своей поверхности. Например, что произойдет, если внешнее поле ортогонально оси симметрии сверхпроводящего цилиндра или же образец представляет собой шар (см. рис. 43.5, *b*)? Из картины линий магнитного поля, параллельных поверхности шара (цилиндра), видно, что напряженность достигает максимального значения  $H_m > H$  вблизи экватора. Связь максимального значения  $H_m$  магнитного поля у поверхности образца с напряженностью  $H$  поля вдали от него можно записать в виде

$$H_m = \frac{H}{1-n}, \quad (43.14)$$

где параметр  $n$  называется *размагничивающим фактором*. Расчеты с помощью обычных методов электростатики дают следующие значения размагничивающего фактора:  $n = 0$  для длинного цилиндра в параллельном его оси магнитном поле,  $n = 1/2$  для длинного цилиндра в поперечном магнитном поле,  $n = 1/3$  для шара и  $n = 1$  для тонкой пластины в магнитном поле, ортогональном ее плоскости. Таким образом, если напряженность магнитного поля вдали от сферического образца равна  $H$ , то вблизи экватора она составляет  $3H/2$  (для определенности мы далее будем говорить о сфере). Пока приложенное поле  $H < 2H_c/3$  (где  $H_c$ , напомним, критическое поле для данной температуры), во всех точках поверхности сферического образца поле будет меньше критического и весь образец будет пребывать в сверхпроводящем состоянии, а магнитное поле будет вытесняться из него. Но что произойдет, если внешнее поле превысит значение  $2H_c/3$ , но останется меньше  $H_c$ ? Можно было бы предположить, что проводник перейдет в нормальное состояние. Но тогда магнитное поле проникнет в его толщу и всюду окажется меньше критического значения  $H_c$ . А это противоречит условию существования сверхпроводящих состояний в слабых полях.



**Рис. 43.12.** Структура промежуточного состояния оловянной сферы при  $T = 2,85$  К и  $H = 0,7H_c$  (заштрихованы нормальные области)

Оказывается, что на самом деле парадокса нет, потому что в образце возникает так называемое *промежуточное состояние*. Образец разобьется на перемежающиеся области (домены) сверхпроводника и нормального металла, и магнитное поле будет проникать в него лишь частично, сквозь области в нормальном состоянии. Структура промежуточного состояния определяется геометрией образца и может представлять собой замысловатую конфигурацию (рис. 43.12 из книги<sup>3</sup>). Характерные размеры доменов в промежуточном состоянии составляют десятые доли миллиметра, что много больше диаметра вихрей Абрикосова. Заметим, что тонкая пластина в поперечном поле переходит в промежуточное состояние уже при бесконечно малом значении внешнего поля  $H$ . Доменная структура в промежуточном состоянии сверхпроводника была впервые экспериментально обнаружена советскими учеными А. Г. Мешковским и А. И. Шальниковым в 1947 г.

## Уравнение Лондонов

Итак, после наблюдения эффекта Мейснера стало ясно, что электродинамика сверхпроводников описывается какими-то дополнительными уравнениями помимо уравнений Максвелла. Их предложили немецкие физики, братья Фриц и Хайнц Лондоны (1935).

Если сопротивление сверхпроводника равно нулю, то носители заряда из сверхпроводящей группы (вспоминаем двухжидкостную модель Гортера—Казимира) свободно ускоряются приложенным электрическим полем. Если масса этих носителей  $m$ , а заряд  $q$ , то второй закон Ньютона имеет вид

$$m \frac{d\vec{v}}{dt} = q\vec{E}. \quad (43.15)$$

В двухжидкостной модели нормальные носители заряда тока не переносят, поэтому полный ток связан только со сверхпроводящими носителями:

$$\vec{j} = nq\vec{v}, \quad (43.16)$$

где  $n$  — концентрация сверхпроводящих носителей заряда. Тогда уравнение (43.15) можно переписать в виде

$$\boxed{\frac{d\vec{j}}{dt} = \frac{nq^2}{m}\vec{E}}. \quad (43.17)$$

В стационарном состоянии  $d\vec{j}/dt = 0$ , откуда следует равенство нулю электрического поля в проводнике:  $\vec{E} = 0$ .

<sup>3</sup> А. Роуз-Инс, Е. Родерик. Введение в физику сверхпроводимости. М.: Мир, 1972. С. 97.

Теперь применяем уравнение Максвелла

$$\operatorname{rot} \vec{E} = -\frac{d\vec{B}}{dt}. \quad (43.18)$$

Выражая напряженность электрического поля через ток согласно (43.17), получаем соотношение

$$\frac{d}{dt} \left( \operatorname{rot} \vec{j} + \frac{nq^2}{m} \vec{B} \right) = 0. \quad (43.19)$$

Величина в круглых скобках сохраняется и равна нулю, поскольку в толще сверхпроводника ток и магнитное поле равны нулю. Таким образом, мы приходим к **уравнению Лондонов**

$$\operatorname{rot} \vec{j} + \frac{nq^2}{m} \vec{B} = 0. \quad (43.20)$$

Уравнения (43.17) и (43.20) лежат в основе электродинамики сверхпроводников, предложенной братьями Лондонами.

Уравнение Лондонов позволяет решить задачу о глубине проникновения магнитного поля в массивный сверхпроводник. Согласно уравнению Максвелла

$$\operatorname{rot} \vec{B} = \mu_0 \vec{j}. \quad (43.21)$$

Применяем к обеим частям (43.21) операцию ротор и используем уравнение (43.20):

$$\operatorname{rot} \operatorname{rot} \vec{B} = \mu_0 \operatorname{rot} \vec{j} = -\frac{nq^2}{\varepsilon_0 m c^2} \vec{B}. \quad (43.22)$$

Используя формулу векторного анализа  $\operatorname{rot} \operatorname{rot} \vec{B} = \operatorname{grad} \operatorname{div} \vec{B} - \nabla^2 \vec{B}$  и уравнение Максвелла  $\operatorname{div} \vec{B} = 0$ , приходим к уравнению для магнитного поля в сверхпроводнике

$$\nabla^2 \vec{B} - \frac{1}{\lambda^2} \vec{B}, \quad \lambda = \sqrt{\frac{\varepsilon_0 m c^2}{nq^2}}. \quad (43.23)$$

Представим, что сверхпроводник занимает полупространство  $z > 0$ , а на границе  $z = 0$  магнитная индукция имеет значение  $B(0)$  и направлена параллельно плоскости проводника. Тогда для направленного так же поля  $B(z)$  в толще проводника получаем из (43.23) уравнение

$$\frac{d^2 B(z)}{dz^2} - \frac{1}{\lambda^2} B(z) = 0, \quad (43.24)$$

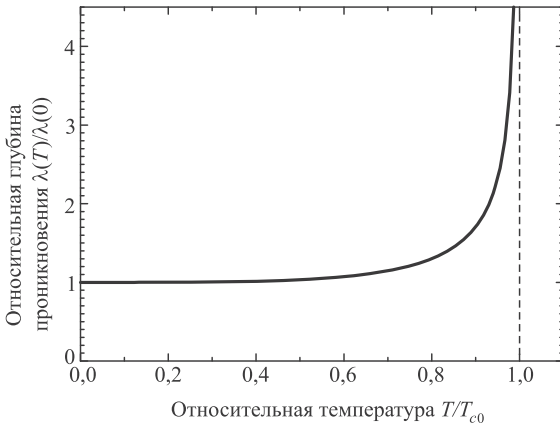
физически осмысленное решение которого имеет вид

$$B(z) = B(0) e^{-z/\lambda}. \quad (43.25)$$

Таким образом, магнитное поле убывает в глубь сверхпроводника по экспоненциальному закону и занимает слой толщиной порядка  $\lambda$ . Параметр  $\lambda$  имеет размерность длины и называется **лондоновской длиной проникновения**. Для численной оценки примем  $n = 10^{28} \text{ м}^{-3}$ , а вместо массы  $m$  и заряда  $q$  возьмем массу и заряд электрона. Имеем

$$\lambda = \sqrt{\frac{8,85 \cdot 10^{-12} \times 9,11 \cdot 10^{-31} \times (3 \cdot 10^8)^2}{10^{28} \times (1,6 \cdot 10^{-19})^2}} \approx 5 \cdot 10^{-8} \text{ м} = 50 \text{ нм}. \quad (43.26)$$





**Рис. 43.13.** Зависимость от температуры лондоновской глубины проникновения магнитного поля в сверхпроводник

В двухжидкостной модели Гортера—Казимира концентрация сверхпроводящих носителей  $n \propto 1 - (T/T_{c0})^4$ , так что глубина проникновения зависит от температуры по закону

$$\lambda(T) = \frac{\lambda(0)}{\sqrt{1 - \left(\frac{T}{T_{c0}}\right)^4}}. \quad (43.27)$$

Эта простая формула описывает экспериментальные данные с точностью до нескольких процентов. При приближении температуры к критическому значению глубина проникновения стремится к бесконечности (рис. 43.13).

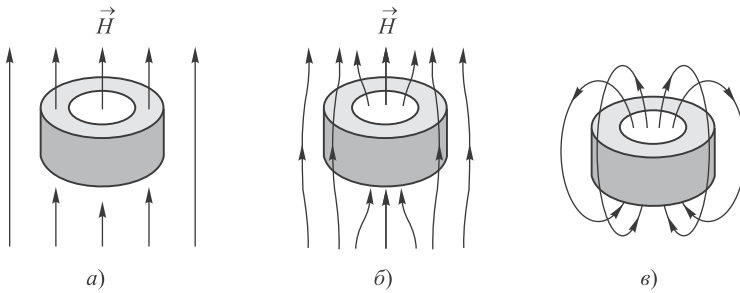
## Квантование магнитного потока

Удивительные явления происходят, если сверхпроводник не является односвязным, как говорят математики, т. е. если не всякий замкнутый путь в нем можно стянуть в точку. Примером может служить обычное кольцо. На рисунке 43.14 изображена последовательность событий, происходящих с кольцом из сверхпроводящего металла. Сначала (а) кольцо в нормальном состоянии помещают в магнитное поле  $\vec{H}$ . Кольцо охлаждают, и оно переходит в сверхпроводящее состояние. Мы уже знаем эффект Мейснера: магнитное поле выталкивается из сверхпроводника. Какая-то часть магнитного потока будет идти при этом сквозь отверстие кольца (б). Теперь внешнее магнитное поле убирают, и линии магнитной индукции, которые шли сквозь отверстие, «замораживаются» (в): исчезновению магнитного потока препятствуют поверхностные сверхпроводящие токи.

Магнитный поток сквозь сверхпроводящее кольцо должен квантоваться, на что впервые обратил внимание Ф. Лондон (1950). Представим себе тонкое сверхпроводящее кольцо радиусом  $R$ , по которому течет ток  $I$ . В главе 22 (см. т. 1), где рассматривается энергия магнитного поля, мы получили формулу (22.32), связывающую энергию  $W_m$ , ток  $I$  и магнитный поток  $\Psi$  в катушке индуктивности:

$$W_m = \frac{I\Psi}{2}. \quad (43.28)$$

Эта формула имеет общий характер, и мы применим ее к сверхпроводящему кольцу. Пусть  $N$  — число носителей заряда в кольце, а  $m$  и  $q$  — их масса и заряд.



**Рис. 43.14.** «Замораживание» магнитного потока сквозь сверхпроводящее кольцо

Тогда ток  $I$  в кольце можно выразить через период  $T$  обращения электронов ( $I = Nq/T$ ), а период — через их скорость  $v = 2\pi R/T$ . В результате получаем из (43.28) энергию магнитного поля (здесь  $q = |e|$ ):

$$W_m = \frac{Nqv\Psi}{4\pi R}. \quad (43.29)$$

В то же время эта энергия получена вследствие того, что электронам придана скорость  $v$ , т. е. сообщена кинетическая энергия

$$W_k = N \frac{mv^2}{2}. \quad (43.30)$$

По закону сохранения энергии  $W_m = W_k$ , и уравнения (43.29) и (43.30) приводят к следующему выражению для магнитного потока:

$$\Psi = \frac{2\pi}{q} mvR. \quad (43.31)$$

Выражение  $L = mvR$  — угловой момент, который, как мы знаем, квантуется:  $L = n\hbar$ ,  $n = 1, 2, 3, \dots$  (значение  $n = 0$ , соответствующее нулевому магнитному потоку, в данной системе невозможно). Иначе, на длине окружности кольца должно укладываться целое число длин волн де Бройля ( $2\pi R = n\lambda = h/(mv)$ ). Мы пришли к **квантованию магнитного потока** сквозь сверхпроводящее кольцо:

$$\Psi = n\Psi_0, \quad \Psi_0 = \frac{h}{q}. \quad (43.32)$$

Здесь возникла новая фундаментальная величина — **квант магнитного потока**  $\Psi_0$ . Масса носителей заряда в ответ не вошла, а их заряд  $q$  Лондон положил равным заряду электрона  $e$ . Эксперименты разных групп подтвердили в 1961 г. заключение о квантовании магнитного потока, но для его кванта (минимальной величины) было найдено значение, вдвое меньшее:

$$\Psi_0 = \frac{h}{2e} = 4,14 \cdot 10^{-15} \text{ Вб}. \quad (43.33)$$

Заметим также, что магнитный поток сквозь вихрь Абрикосова в точности равен одному кванту магнитного потока (43.33). В СИ единицей магнитного потока служит вебер ( $1 \text{ Вб} = 1 \text{ кг} \cdot \text{м}^2 \cdot \text{с}^{-2} \cdot \text{А}^{-1} = 1 \text{ Тл} \cdot \text{м}^2$ ).

Таким образом, *результаты экспериментов указывали, что заряд носителей тока в сверхпроводнике равен удвоенному заряду электрона:  $q = 2e$ .*

## Электрон-фононное взаимодействие и куперовские пары

Как мы видели в предыдущей главе, теплоемкость нормального металла при низких температурах имеет вид

$$C = A \left( \frac{T}{\theta_D} \right)^3 + \gamma T, \quad (43.34)$$

где первое слагаемое — это уже знакомый нам вклад кристаллической решетки, описываемый теорией Дебая, а второе слагаемое — вклад электронного газа. Эксперименты показывают, что в сверхпроводнике при очень низких температурах теплоемкость ведет себя иначе:

$$C = A \left( \frac{T}{\theta_D} \right)^3 + ae^{-b/(k_B T)}. \quad (43.35)$$

Первое слагаемое не меняется, и даже коэффициент  $A$  остается тем же самым. Это значит, что по-прежнему на месте вклад кристаллической решетки. Драматическое изменение второго слагаемого (экспонента вместо линейной функции) показывает, что сверхпроводимость связана с какими-то существенными изменениями поведения электронов проводимости.

В то же время в 1950 г. были проведены эксперименты, в которых изучалось сверхпроводящее состояние не естественной ртути  $^{202}\text{Hg}$ , а ее изотопа  $^{198}\text{Hg}$ . Оказалось, что критическая температура обратно пропорциональна массе атома ртути:  $T_{c0} \propto 1/\sqrt{m}$ . Но такую же зависимость имеет и дебаевская температура  $\theta_D$ . Действительно, согласно уравнению (40.30), дебаевская температура пропорциональна максимальной частоте  $\omega_{\max}$  колебаний атомов кристаллической решетки. В свою очередь, в разделе о фононах мы видели, что кристаллическую решетку можно представить как набор осцилляторов массой  $m$ , связанных «пружинами» жесткостью  $k$ . Тогда масштаб частот колебаний решетки задается частотой колебаний отдельного осциллятора:  $\omega_0 = \sqrt{k/m}$ . Зависимость критической температуры перехода вещества в сверхпроводящее состояние от массы атомов называется *изотопическим эффектом*.

Что же у нас выходит? Из измерений теплоемкости и прямого рентгенографического исследования кристаллов следует, что кристаллическая решетка не меняется при переходе образца в сверхпроводящее состояние. А изотопический эффект показывает, что она играет важную роль в явлении сверхпроводимости. В 1950 г. немецко-британский физик Г. Фрëлих заметил, что электроны могут взаимодействовать с фононами — квантами колебаний кристаллической решетки. Согласно Фрëлиху, электрон может испустить виртуальный фонон, который сразу же будет поглощен другим электроном. Этот процесс подобен обмену фотонами между заряженными частицами в квантовой электродинамике и другим подобным процессам. Обмен квантами поля — основа описания всех фундаментальных взаимодействий, существующих в природе. Фрëлих показал, что обмен фононами может привести к слабому притяжению электронов между собой. На классическом языке взаимодействие электронов посредством обмена фононами можно представить себе следующим образом. Пролетая через кристалл, электрон поляризует среду, возбуждает колебания решетки. Притягивающиеся к нему положительно заряженные ионы кристаллической решетки

создают область поляризации с избытком положительного заряда. Если область поляризации не успевает рассосаться к моменту пролета второго электрона, она также воздействует и на него. Притягиваясь к одной и той же положительно заряженной области поляризации, электроны притягиваются друг к другу.

Гипотеза о связи сверхпроводимости с электрон-фононным взаимодействием позволила Фрёлыху предсказать изотопический эффект до того, как он был открыт экспериментально. Стал понятен также тот факт, что сверхпроводники являются плохими проводниками в нормальном состоянии. Например, у свинца довольно высокая критическая температура (7,2 К), так что в нем электрон-фононное взаимодействие должно быть сильным. Это значит, что длина свободного пробега, обусловленная взаимодействием с кристаллической решеткой, невелика. Соответственно свинец плохо проводит ток в нормальном состоянии при комнатной температуре (удельное сопротивление  $\rho_{\text{Pb}} = 0,22 \text{ Ом} \cdot \text{мм}^2/\text{м}$ ). Наоборот, золото и серебро являются хорошими проводниками в нормальных условиях ( $\rho_{\text{Au}} = 0,023 \text{ Ом} \cdot \text{мм}^2/\text{м}$ ,  $\rho_{\text{Ag}} = 0,015 \text{ Ом} \cdot \text{мм}^2/\text{м}$ ). Это означает, что велика длина свободного электронов, т. е. они слабо взаимодействуют с кристаллической решеткой. Соответственно некоторые хорошие проводники (эти и другие — литий, калий, натрий, медь) не переходят в сверхпроводящее состояние при самых низких достигнутых температурах, а критическая температура других хороших проводников очень низка (например, для алюминия  $T_{c0} = 1,18 \text{ К}$ , а для бериллия  $T_{c0} = 0,024 \text{ К}$ ).

Следующий шаг сделал американский физик Л. Купер. В 1956 г. он рассмотрел задачу о добавлении к металлу при абсолютном нуле температуры двух электронов. В силу принципа Паули они должны заполнить состояния вне сферы Ферми (т. е. с импульсами  $p > p_F$ ) и кинетическая энергия системы должна повыситься, как минимум, на величину  $2E_F$ . При отсутствии взаимодействия между электронами повысится и полная энергия системы. Но Купер показал, что если между электронами существует слабое притяжение, то они могут образовать связанное состояние (т. е. состояние с отрицательной потенциальной энергией). Понижение потенциальной энергии за счет взаимодействия превышает разность между кинетической энергией пары и величиной  $2E_F$ . Поэтому полная энергия системы **понижится**. Квантовомеханические расчеты показывают, что максимальное понижение энергии соответствует случаю, когда спаренные электроны имеют равные и противоположно направленные импульсы, а также и противоположно направленные спины, так что полный импульс и полный спин пары равны нулю. Такие состояния получили название **куперовских пар**.

## Микроскопическая теория сверхпроводимости (БКШ)

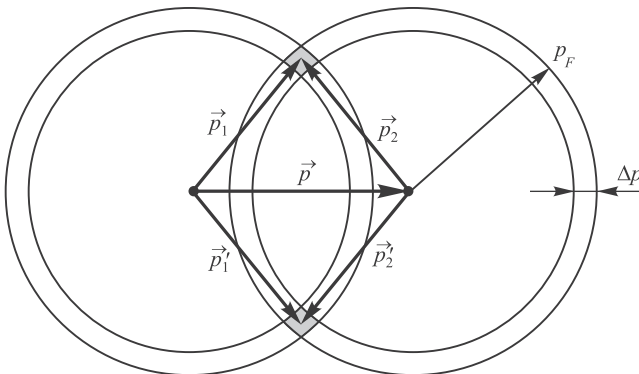
Микроскопическая теория сверхпроводимости (БКШ-теория) была создана американскими учеными Дж. Бардиным, Л. Купером и Дж. Шриффером (1957) на основе представлений о куперовских парах. Микроскопическая теория сверхпроводимости была усовершенствована Н. Н. Боголюбовым (1958), который использовал изящный математический метод, ранее развитый им в теории сверхтекучести. Купер рассматривал добавление в систему двух электронов. В БКШ-теории эти результаты были применены к случаю, когда спариваются два электрона

с противоположно направленными спинами, находящиеся в бесконечно малой окрестности сферы Ферми и имеющие равные по модулю, но противоположно направленные импульсы. Понижение потенциальной энергии пары приводит к энергетически более выгодному состоянию, чем ступенчатое распределение Ферми для нормального металла при нуле температуры. То, что может произойти с одной парой электронов, может случиться и с другой: следующее спаривание электронов еще более понизит энергию системы. При спаривании электроны из заполненных состояний должны переходить в незаполненные. По мере того как все большее число куперовских пар заполняет состояния вне сферы Ферми ( $p > p_F$ ), находить незанятые состояния становится все труднее, и пары располагаются все выше над сферой Ферми. В какой-то момент понижение потенциальной энергии уже окажется недостаточным для компенсации роста кинетической энергии пары, и дальнейшее понижение энергии системы за счет спаривания электронов окажется невозможным. Можно оценить толщину слоя вблизи поверхности Ферми, в котором электроны могут спариваться. При обмене фононами спаренный электрон может выйти за пределы сферы Ферми, приобретя энергию  $E \approx E_F + h\nu$ , где  $\nu$  — характерная частота фонона в данной кристаллической решетке, равная примерно половине дебаевской частоты. Получаем тогда оценку для импульса этого электрона:

$$p = \sqrt{2m_e E} = \sqrt{2m_e(E_F + h\nu)} = \sqrt{p_F^2 + 2m_e h\nu} \approx p_F + \frac{m_e h\nu}{p_F}. \quad (43.36)$$

Таким образом, в куперовские пары объединятся не все электроны, а лишь находящиеся в тонком слое толщиной  $\Delta p \approx m_e h\nu/p_F$ .

Равенство нулю полного импульса куперовской пары иллюстрируется на рис. 43.15. Два электрона, взаимодействующие друг с другом посредством обмена фононами, переходят из состояния с импульсами  $\vec{p}_1$  и  $\vec{p}_2$  в состояние с импульсами  $\vec{p}'_1$  и  $\vec{p}'_2$ . Для образования куперовской пары необходимо, во-первых, чтобы все эти импульсы лежали в тонком слое толщиной  $\Delta p$  вблизи поверхности Ферми, т. е. сферы радиусом  $p_F$ . Во-вторых, вследствие сохранения полного импульса должно выполняться соотношение  $\vec{p}_1 + \vec{p}_2 = \vec{p}'_1 + \vec{p}'_2 = \vec{p}$ . Величина  $p$  равна расстоянию между центрами сфер на рис. 43.15, а число пар, импульсы которых



**Рис. 43.15.** Иллюстрация равенства нулю полного импульса куперовской пары

удовлетворяют обоим указанным условиям, пропорционально объему заштрихованных областей. Этот объем максимален при  $\vec{p} = 0$ .

Спин электронов равен половине, поэтому куперовские пары заведомо имеют целый спин  $s$ , который может быть равен нулю или единице. Однако состояние с  $s = 1$  неустойчиво и полный спин куперовской пары равен нулю. Поэтому куперовские пары могут рассматриваться как частицы с нулевым спином. Являясь бозонами, они могут находиться одним и тем же состоянии с наименьшей энергией, т. е. в состоянии бозе—эйнштейновского конденсата. Но это, конечно, очень необычные частицы. Расчеты дают, что расстояние между спаренными электронами (диаметр куперовской пары) довольно велико — примерно 1 мкм. Отсюда следует, что объем куперовской пары по порядку величины равен  $V_C \sim 10^{-18} \text{ м}^3$ . Типичная концентрация электронов в металле  $n \sim 10^{27} \text{ 1/м}^3$ , и такой же должна быть концентрация куперовских пар при температуре, близкой к абсолютному нулю. Стало быть, в объеме одной куперовской пары помещается около миллиарда ( $nV_C \sim 10^9$ ) других куперовских пар! Иначе говоря, среднее расстояние между центрами масс куперовских пар в тысячу раз меньше размера самих пар и равно примерно 1 нм. Таким образом, куперовскую пару никоим образом невозможно представить как квазимолекулу. Спаренные электроны близки друг к другу *в импульсном пространстве*, а в обычном координатном пространстве пары проникают друг в друга (их волновые функции сильно перекрываются, если говорить на языке квантовой механики). Трудно представить себе, что электрон «чувствует» своего партнера, несмотря на то что между ними мельтешит столько назойливых соседей из других куперовских пар, но именно так устроена природа.

Чрезвычайно важно, что основное состояние (конденсат) отделено энергетической щелью  $E_g$  от следующего, возбужденного состояния. Иначе говоря, чтобы разорвать одну куперовскую пару и перевести систему в первое возбужденное состояние, надо придать ей некоторую минимальную энергию  $E_g$ , а меньшие порции энергии просто не будут восприняты системой. Нечто подобное мы уже видели для сверхтекучего состояния, а еще ранее для теплоемкости («замораживание» вращательных и колебательных степеней свободы, теория теплоемкости Эйнштейна). Поэтому можно сказать, что сверхпроводимость — это сверхтекучесть газа куперовских пар. Они и представляют собой сверхтекучий компонент, угаданный в ранних двухжидкостных теориях сверхпроводимости. БКШ-теория объясняет также упомянутый выше двойной заряд носителей тока в сверхпроводнике, обнаруженный при наблюдениях квантования магнитного потока.

Согласно теории БКШ, энергия, необходимая для разрыва куперовской пары и образования двух свободных электронов с импульсами  $p_1$  и  $p_2$ , равна:

$$E = E_1 + E_2, \quad E_i = \left[ (E_i - E_F)^2 + \Delta^2 \right]^{1/2} \quad (i = 1, 2), \quad (43.37)$$

где  $E_i = p_i^2/(2m)$ , а величина  $\Delta$  равна половине энергетической щели. Действительно, минимальная энергия, требуемая для разрыва куперовской пары, получается, когда образующиеся свободные электроны оказываются на поверхности сферы Ферми:  $p_i = p_F$ , т. е.  $E_i = E_F$ . Тогда из (43.37) следует, что в спектре возбуждения сверхпроводника имеется энергетическая щель шириной

$E_g = \min E = 2\Delta$ . Поскольку энергия  $E$  в (43.37) равна сумме двух слагаемых, каждое из которых зависит от состояния лишь одного электрона, можно считать, что величина  $E_i$  — это вклад в энергию сверхпроводника  $i$ -го электрона, возникающего при разрыве пары. Поэтому схему энергетических уровней сверхпроводника можно представить себе как низший уровень энергии, общий для всех сконденсированных куперовских пар, и начинающийся на расстоянии  $\Delta$  над ним континуум разрешенных энергетических уровней — возбужденных состояний свободных электронов.

Теория БКШ предсказывает простую связь энергетической щели и критической температуры:  $E_g = 3,53k_B T_{c0}$ . Наличие энергетической щели сразу объясняет поведение теплоемкости электронного газа. Действительно, энергия  $E_g$  — это минимальная энергия, разрушающая куперовскую пару, и ее передача системе порождает пару свободных электронов. Тогда на каждый электрон приходится половинная энергия  $\Delta = E_g/2$ . Из статистических соображений можно утверждать тогда, что число электронов вне основного состояния пропорционально величине  $e^{-E_g/(2k_B T)} = e^{-1,75T_{c0}/T}$ . Тепловая энергия, поглощаемая при возбуждении, пропорциональна  $E_g e^{-1,75T_{c0}/T}$ . Производная этой величины по температуре дает теплоемкость электронного газа в сверхпроводящем состоянии:

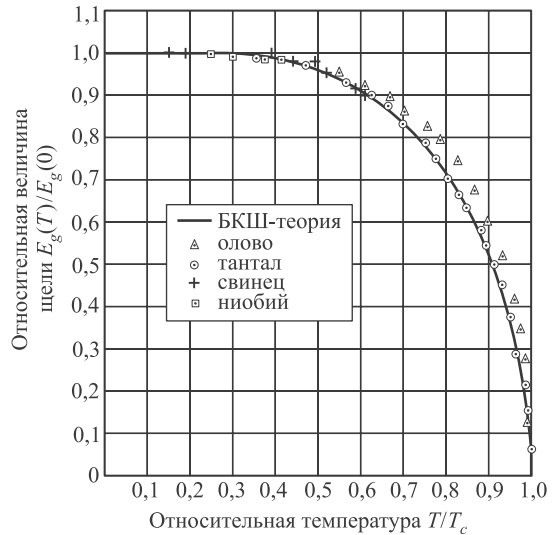
$$C_{эл} = \text{const} \left( \frac{T_{c0}}{T} \right)^2 \exp \left( -1,75 \frac{T_{c0}}{T} \right). \quad (43.38)$$

При низких температурах предэкспоненциальный множитель  $1/T^2$  меняется гораздо медленнее экспоненты, так что этот закон практически не отличается от феноменологического соотношения (43.35). Заметим, что выражение (43.38) аналогично формуле (40.12) для теплоемкости в теории Эйнштейна, где также имеется щель между основным и возбужденным уровнями энергии.

При повышении температуры число куперовских пар уменьшается. Во-первых, тепловое возбуждение способствует разрыву пар и перескоку электронов через энергетическую щель. Во-вторых, сама величина щели уменьшается, так как становится меньше число актов взаимодействия электронов с фононами. На классическом языке тепловое возбуждение быстрее разрушает область поляризации с избытком положительного заряда, которая и приводит к притяжению электронов. Изменение величины энергетической щели показано на рис. 43.16. При критической температуре щель исчезает и вместе с ней исчезают сверхпроводящие свойства.

Из сказанного ясно, какую важную роль играет энергетическая щель. В системе фермионов, как мы видели, она ответственна за различие свойств металлов, диэлектриков и полупроводников, а в системе бозонов — за явления сверхтекучести и сверхпроводимости.

В 1972 г. Бардин, Купер и Шриффер получили Нобелевскую премию «за создание теории сверхпроводимости, обычно называемой БКШ-теорией». Тем самым Дж. Бардин стал дважды нобелевским лауреатом по физике (первую премию он получил в 1956 г. «за исследования полупроводников и открытие транзисторного эффекта»). Но было бы самонадеянно утверждать, что ученые уже все поняли в механизме сверхпроводимости. Исследования этого удивительного



**Рис. 43.16.** Экспериментальные данные о зависимости энергетической щели от температуры для некоторых сверхпроводников (из работы *P. Townsend, J. Sutton. Phys. Rev.* **128** (2), 591, 1962)

явления, в особенности высокотемпературных сверхпроводников, продолжают, и нет сомнений, что эта область физики еще поразит нас неожиданностями.

### Туннелирование электронов при контакте сверхпроводников

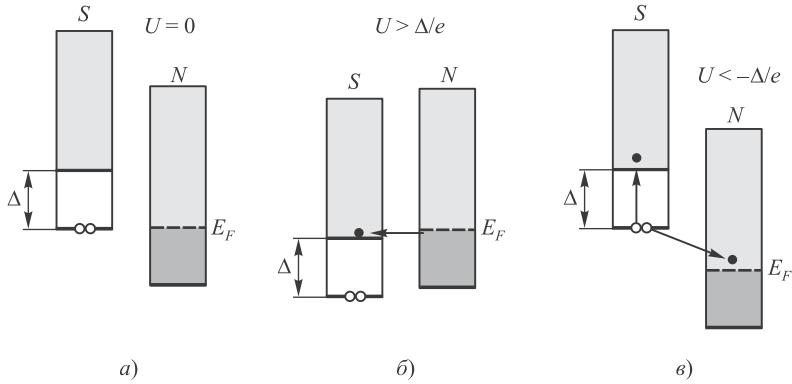
В этом разделе мы поговорим о процессах туннелирования электронов через контакт сверхпроводника с нормальным металлом или другим сверхпроводником. Здесь существует два различных механизма. Один из них — туннелирование свободных электронов — напоминает процессы в туннельном диоде (хотя в некоторых чертах существенно от них отличается). Им мы для начала и займемся.

#### Туннелирование между сверхпроводником и нормальным металлом

При нулевой температуре все конденсированные куперовские пары сверхпроводника располагаются на низшем энергетическом уровне. Выше этого уровня на величину  $\Delta$  начинается континуум разрешенных уровней энергии электронов. В нормальном же металле электроны заполняют все уровни энергии ниже уровня энергии Ферми  $E_F$ . При контакте сверхпроводника и нормального металла при нулевой разности потенциалов между ними уровень энергии куперовских пар соответствует по высоте уровню энергии Ферми нормального металла (рис. 43.17, *a*). Это состояние термодинамического равновесия, когда ток через такой контакт не идет.

Если подать на сверхпроводник положительное напряжение  $U$ , энергетические уровни сверхпроводника опустятся (заряд электрона  $-e$  отрицателен). Однако при небольших напряжениях процесс туннелирования невозможен и ток по-прежнему не идет. При повышении напряжения до величины  $U = \Delta/e$  нижняя граница континуума незаполненных уровней опускается ниже уровня Ферми. Тогда свободные уровни сверхпроводника оказываются на высоте заполненных





**Рис. 43.17.** Энергетическая диаграмма для контакта сверхпроводника ( $S$ ) и нормального металла ( $N$ ) при  $T = 0$  К и различных разностях потенциала между ними; светлым тоном показаны области с разрешенными свободными уровнями энергии, темным тоном — области в нормальном металле с заполненными уровнями ниже уровня энергии Ферми  $E_F$ ; соединенные светлые кружочки обозначают куперовские пары в сверхпроводнике

уровней нормального металла и электроны туннелируют из нормального металла в сверхпроводник (рис. 43.17, б). С ростом поданного на сверхпроводник напряжения растет число туннелирующих электронов и, следовательно, увеличивается туннельный ток.

Если к сверхпроводнику приложить отрицательное напряжение, то его энергетические уровни сместятся вверх, описанный процесс туннелирования электронов будет невозможен и ток через контакт идти не будет. Однако с ростом напряжения становится возможным другой процесс — расщепление куперовских пар. Напомним, что мы рассматриваем *изоэнергетические процессы*, в которых отсутствует тепловая энергия возбуждения электронов. На энергетической диаграмме одиночный электрон может перейти из одного металла в другой лишь «по горизонтали», так как его энергия должна оставаться постоянной. При расщеплении же куперовской пары сохранение полной энергии двух электронов возможно за счет того, что один из них приобретет энергию  $\Delta$ , необходимую для перехода на свободный энергетический уровень в континууме, а другой потеряет точно такую же энергию. Чтобы этот второй электрон попал на свободный уровень, лежащий в нормальном металле выше уровня Ферми, этот последний должен лежать ниже уровня куперовских пар по крайней мере на величину  $\Delta$  (рис. 43.17, в). Стало быть, такой процесс становится возможным по достижении напряжения  $U = -\Delta/e$ . При дальнейшем росте абсолютной величины отрицательного напряжения на сверхпроводнике число расщепляющихся пар растет и соответственно возрастает обратный ток через контакт.

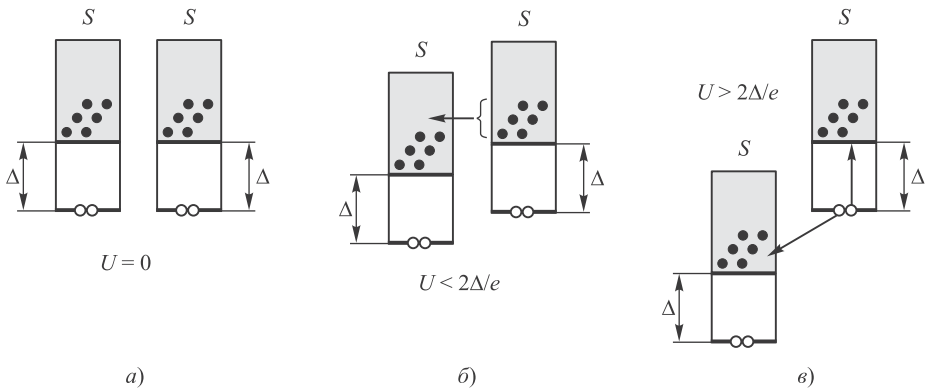
В такого рода процессах можно непосредственно измерить величину  $\Delta$  данного сверхпроводника по напряжению, при котором неожиданно возникает туннельный ток между сверхпроводником и любым нормальным металлом.

При температурах выше абсолютного нуля часть электронов нормального металла располагается на энергетической диаграмме выше уровня Ферми и при подаче положительного напряжения они могут туннелировать в свободные состояния в энергетическом континууме сверхпроводника, так что при  $T \neq 0$  через контакт будет идти небольшой ток даже при напряжениях  $0 < U < \Delta/e$ . Ниже уровня Ферми образуются незаполненные состояния, куда может туннелировать одна из частиц куперовской пары при отрицательном напряжении, так что небольшой ток пойдет и при напряжениях  $-\Delta/e < U < 0$ . Тем не менее и при ненулевых температурах при достижении напряжений  $U = \pm\Delta/e$  будет наблюдаться резкий рост туннельного тока.

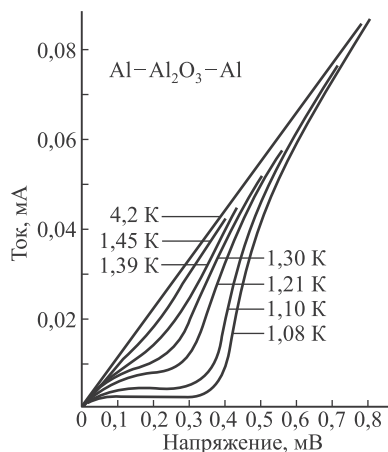
### Туннелирование электронов между двумя сверхпроводниками

Если в контакте находятся два одинаковых сверхпроводника, то при температурах выше абсолютного нуля в континууме разрешенных энергетических уровней всегда имеется некоторое количество свободных электронов, занимающих вследствие теплового возбуждения полосу шириной  $k_B T$  (рис. 43.18, *a*). Поскольку сверхпроводники находятся при температурах жидкого гелия порядка 1–2 К, ширина этой полосы равна примерно  $10^{-4}$  эВ. Если подать на левый сверхпроводник напряжение  $U > 10^{-4}$  В, его энергетические уровни опустятся, и напротив тепловых электронов правого проводника окажутся свободные уровни в левом. Возникнет поток электронов справа налево, т. е. ток через контакт (рис. 43.18, *б*). Но этот ток невелик, так как мало число возбужденных электронов и, что характерно, ток остается почти постоянным, не увеличивается с ростом напряжения, так как число возбужденных электронов от этого не возрастает.

Однако по достижении напряжения  $U = 2\Delta/e$  становится возможным процесс расщепления куперовских пар в правом сверхпроводнике: один из электронов



**Рис. 43.18.** Энергетическая диаграмма для контакта двух одинаковых сверхпроводников при  $T > 0$  и различных разностях потенциала между ними; вследствие теплового возбуждения выше уровня куперовских пар имеется некоторое количество свободных электронов (показаны черными кружочками)



**Рис. 43.19.** Вольт-амперная характеристика Al-Al-контакта при разных температурах; температура перехода алюминия в сверхпроводящее состояние  $T_{c0} = 1,18$  К

пары туннелирует налево, заполняя низшее доступное состояние в энергетическом континууме и теряя энергию  $\Delta$ , а другой электрон, приобретая такую же энергию, переходит на свободный энергетический уровень в континууме правого сверхпроводника (рис. 43.18, в). Полная энергия при этом сохраняется: процесс является изоэнергетическим. При увеличении напряжения резко возрастает число расщепляющихся куперовских пар, потому что туннелирующим электронам становятся доступными все более высокие энергетические уровни. Соответственно, туннельный ток быстро растет с напряжением, достигая величин, обычных для омического сопротивления нормальных металлов.

На рисунке 43.19 из книги Д. Р. Гилли и Дж. Гилли (см. сноску на с. 451) показана вольт-амперная характеристика контакта двух сверхпроводящих образцов алюминия, разделенных тонкой пленкой окиси, измеренная в одной из работ первооткрывателя туннелирования в сверхпроводниках Айвара Джайевера (Ивара Гиевера), американского физика норвежского происхождения. В 1973 г. Л. Эсаки и А. Джайевер получили половину Нобелевской премии «за экспериментальные открытия туннельных явлений в полупроводниках и сверхпроводниках соответственно».

## Эффекты Джозефсона

Вторую половину Нобелевской премии в том же 1973 г. получил английский физик-теоретик Б. Джозефсон «за теоретическое предсказание свойств тока сверхпроводимости, проходящего через туннельный барьер, в частности явлений, обычно называемых эффектом Джозефсона». В 1962 г., в возрасте 22 лет, будучи никому не известным аспирантом Кембриджского университета, он предсказал два новых удивительных явления, возникающие при пропускании тока от внешнего источника через контакт двух сверхпроводников, разделенных очень узким (порядка 1 нм) зазором из нормального металла или даже диэлектрика (теперь такую конструкцию называют *контактом Джозефсона*).

**Стационарный эффект Джозефсона** заключается в протекании тока через такой контакт, даже если «прослойка» между сверхпроводниками сама по

себе сверхпроводником не является или даже вообще не проводит электрический ток. Этот ток через контакт не создает на нем падения напряжения, т. е. является «сверхтоком», но по величине он гораздо слабее «сверхтоков» в массивных сверхпроводниках. Такой ток идет лишь при условии, что его величина не превышает некоторого критического значения  $I_0$ .

**Нестационарный эффект Джозефсона** состоит в излучении высокочастотных электромагнитных волн из области контакта сверхпроводников. Оказывается, что при увеличении силы внешнего тока на контакте Джозефсона появляется падение напряжения  $U$ , которое, помимо постоянной составляющей  $\bar{U}$ , имеет и осциллирующую часть, приводящую к излучению. Теоретические предсказания Джозефсона вскоре были подтверждены экспериментально. Сейчас эффекты Джозефсона находят широкое практическое применение.

В эффектах Джозефсона проявляется другой механизм возникновения туннельного тока, нежели туннелирование электронов-одиночек: здесь через потенциальный барьер, созданный узким зазором, туннелируют куперовские пары целиком. Иными словами, после прохождения зазора между сверхпроводниками туннелирующие электроны остаются спаренными. Мы уже знаем, что все куперовские пары в сверхпроводнике описываются единой волновой функцией, являющейся комплексной величиной, квадрат амплитуды которой равен концентрации частиц в сверхпроводящем компоненте. Фазы же волновых функций в двух разделенных сверхпроводниках могут принимать произвольные значения, но при очень малой ширине зазора волновые функции конденсата в двух сверхпроводниках сильно перекрываются и возникают эффекты интерференции.

В дальнейшем изложении мы следуем Р. Фейнману<sup>4</sup>, предложившему простой и изящный вывод уравнений, описывающих эффекты Джозефсона. Пусть  $\psi_1$  — волновая функция конденсата по одну сторону барьера, а  $\psi_2$  — соответствующая волновая функция конденсата по другую его сторону. Обозначим  $E_1$  и  $E_2$  энергии куперовских пар в бозе-конденсате, а символом  $K$  — постоянную, характеризующую переход спаренных электронов из одного материала в другой. Простейшие уравнения Шрёдингера для этих состояний напоминают уравнения колебаний связанных осцилляторов или уравнения, описывающие двухуровневую среду для генерации лазерного излучения:

$$\begin{aligned} i\hbar \frac{\partial \psi_1}{\partial t} &= E_1 \psi_1 + K \psi_2, \\ i\hbar \frac{\partial \psi_2}{\partial t} &= E_2 \psi_2 + K \psi_1. \end{aligned} \quad (43.39)$$

Если бы постоянная  $K$  была равна нулю, связи между конденсатом в двух сверхпроводниках не было бы и уравнения (43.39) описывали бы гармонические колебания — волны де Бройля куперовских пар, независимо распространяющиеся в каждом сверхпроводнике. Уравнения (43.39) описывают так называемую *слабую связь*, так как они учитывают лишь линейную зависимость изменения волновой функции в одном сверхпроводнике от волновой функции в другом. Подставляем в эти уравнения волновые функции в сверхпроводниках по обе стороны зазора в виде

$$\begin{aligned} \psi_1 &= \sqrt{n_1} e^{i\theta_1}, \\ \psi_2 &= \sqrt{n_2} e^{i\theta_2}. \end{aligned} \quad (43.40)$$

<sup>4</sup> Р. Фейнман, Р. Лейтон, М. Сэндс. Фейнмановские лекции по физике. Вып. 8–9. М.: Мир, 1978. С. 512.

Здесь  $n_1$  и  $n_2$  — концентрации куперовских пар в двух сверхпроводниках<sup>5</sup>. Подставляя (43.40) в (43.39), мы можем переписать уравнения для волновых функций в терминах уравнений для фаз и концентраций частиц:

$$\begin{aligned} i\hbar \frac{\dot{n}_1}{2\sqrt{n_1}} e^{i\theta_1} - \hbar \dot{\theta}_1 \sqrt{n_1} e^{i\theta_1} &= E_1 \sqrt{n_1} e^{i\theta_1} + K \sqrt{n_2} e^{i\theta_2}, \\ i\hbar \frac{\dot{n}_2}{2\sqrt{n_2}} e^{i\theta_2} - \hbar \dot{\theta}_2 \sqrt{n_2} e^{i\theta_2} &= E_2 \sqrt{n_2} e^{i\theta_2} + K \sqrt{n_1} e^{i\theta_1}, \end{aligned} \quad (43.41)$$

где для упрощения уравнений точками обозначены производные по времени:  $\dot{n}_1 = \partial n_1 / \partial t$  и т. д.

Приравнявая мнимые части в левых и правых сторонах этих уравнений, получаем уравнения для изменения концентрации частиц на границах контакта:

$$\begin{aligned} \dot{n}_1 &= \frac{2K}{\hbar} \sqrt{n_1 n_2} \sin \delta, \\ \dot{n}_2 &= -\frac{2K}{\hbar} \sqrt{n_1 n_2} \sin \delta, \end{aligned} \quad (43.42)$$

где символом  $\delta$  мы обозначили скачок фазы волновой функции на границе контакта. Сверхпроводящий ток через контакт пропорционален изменению концентрации носителей заряда на границах контакта, поэтому для него можно написать выражение, аналогичное (43.42):

$$\boxed{I_s = I_0 \sin \delta.} \quad (43.43)$$

Здесь  $I_0$  — максимально возможное значение тока через контакт, пропорциональное коэффициенту связи  $K$  и концентрации носителей заряда в сверхпроводниках.

Приравнявая действительные части в обеих сторонах уравнений (43.41), получаем уравнения для эволюции фаз волновых функций:

$$\begin{aligned} \dot{\theta}_1 &= -E_1 - K \sqrt{\frac{n_2}{n_1}} \cos \delta, \\ \dot{\theta}_2 &= -E_2 - K \sqrt{\frac{n_1}{n_2}} \cos \delta. \end{aligned} \quad (43.44)$$

Вычитая из второго уравнение первое и учитывая, что концентрации носителей заряда в обоих сверхпроводниках практически равны ( $n_1 \approx n_2$ ), получаем уравнение для разности фаз  $\delta = \theta_2 - \theta_1$ :

$$\boxed{\dot{\delta} = -\frac{E_2 - E_1}{\hbar}.} \quad (43.45)$$

Уравнения (43.43) и (43.45) являются основными уравнениями, описывающими эффекты Джозефсона.

## Стационарный эффект Джозефсона

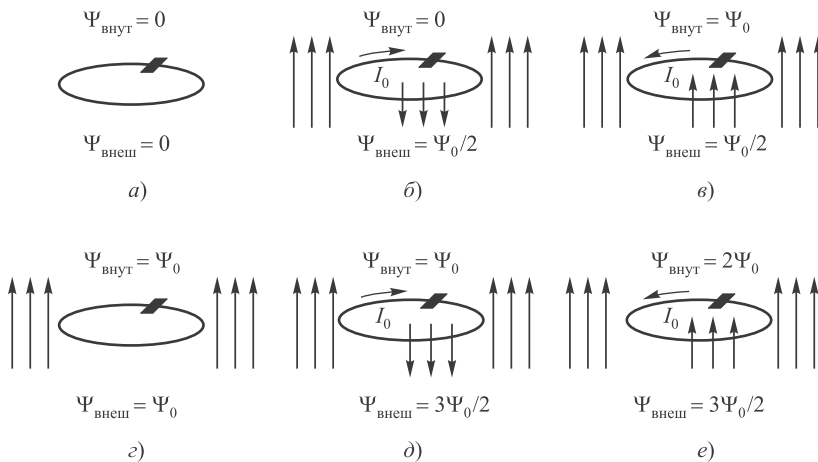
Согласно уравнению (43.43), величина туннельного тока куперовских пар определяется скачком фазы волновой функции по обе стороны зазора между сверхпроводниками, причем этот ток не может превышать некоторого критического

<sup>5</sup>Напомним, что квадрат модуля волновой функции есть плотность вероятности нахождения частицы в данной точке пространства, т. е. величина  $|\psi|^2$  должна быть пропорциональна концентрации частиц, откуда и появились квадратные корни в (43.40).

значения  $I_0$ , которое достигается при разности фаз  $\delta = \pi/2$ . Если подсоединить к туннельному контакту внешний контур и пропускать через него ток  $0 \leq I_s \leq I_0$ , то вольт-амперная характеристика будет представлять собой вертикальный отрезок, расположенный в точке  $U = 0$ . Иными словами, при токе, меньшем критического, падения напряжения на контакте нет, т. е. его сопротивление равно нулю. Если сверхпроводники по обе стороны контакта Джозефсона одинаковы, то в уравнении (43.45) энергии пар равны ( $E_1 = E_2$ ) и скачок фазы  $\delta$  не зависит от времени, т. е. через контакт идет **постоянный ток Джозефсона**. Типичное значение критического тока примерно равно 1 мВ.

Критический ток  $I_0$  пропорционален константе связи  $K$  сверхпроводников, а значит, при контакте двух одинаковых сверхпроводников этот ток пропорционален концентрации  $n_s = n_1 = n_2$  сверхтекучих носителей заряда (куперовских пар). Чем выше температура сверхпроводника, тем уже энергетическая щель и тем меньше концентрация куперовских пар и значение критического тока. В модели Гортера—Казимира  $n_s \propto 1 - (T/T_{c0})^4$ , т. е. вблизи критической температуры  $T_{c0}$ , концентрация носителей заряда, а вместе с ней и критический ток в контакте двух одинаковых сверхпроводников, зависят от температуры по закону  $I_0 \propto n_s \propto T_{c0} - T$ .

Представим теперь, что сверхпроводник свернут в замкнутый контур, разорванный в одном месте контактом Джозефсона (рис. 43.20, а). При отсутствии внешнего поля магнитный поток через контур равен нулю. Если включить слабое внешнее магнитное поле с потоком  $\Psi_{\text{внеш}}$ , то нулевое значение потока через сверхпроводящий контур будет «заморожено», поток по-прежнему будет равен нулю. Это означает, что в контуре возник индукционный ток, магнитный поток



**Рис. 43.20.** Сверхпроводящий контур с контактом Джозефсона в магнитном поле: длинные параллельные стрелки вне контура показывают направление внешнего магнитного потока  $\Psi_{\text{внеш}}$ , короткие стрелки внутри контура показывают направление магнитного потока  $\Psi_{\text{инд}}$ , создаваемого током в контуре, а полный поток через контур равен алгебраической сумме  $\Psi_{\text{внут}} = \Psi_{\text{внеш}} + \Psi_{\text{инд}}$

от которого  $\Psi_{\text{инд}} = -\Psi_{\text{внеш}}$  полностью компенсирует поток внешнего магнитного поля, так что полный поток через контур равен нулю:  $\Psi_{\text{внут}} = \Psi_{\text{внеш}} + \Psi_{\text{инд}}$ . Повышение напряженности внешнего магнитного поля ведет к росту индукционного тока.

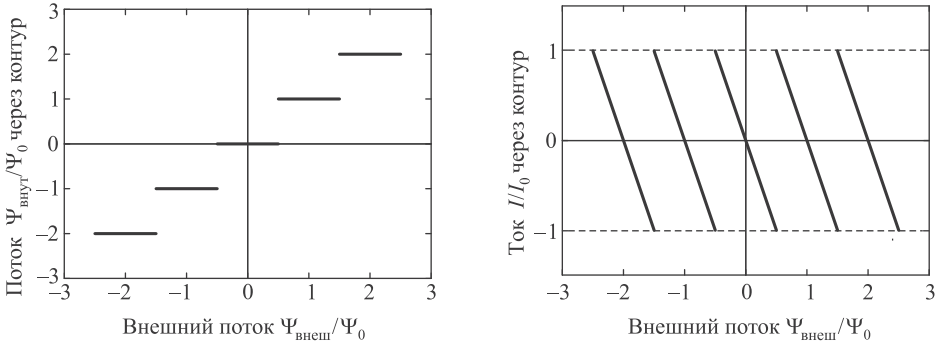
Далее мы следуем наглядным рассуждениям из учебного пособия<sup>6</sup> и предполагаем для простоты, что критический ток  $I_0$  достигается при потоке внешнего поля  $\Psi_{\text{внеш}} = \Psi_0/2$ , равном половине кванта магнитного потока  $\Psi_0$ , определенно-го выше уравнением (43.32). Этого всегда можно добиться, но принципиального значения для окончательных выводов не имеет. На рисунке 43.20, б показана ситуация, когда внешнее поле почти достигло этого значения, а ток почти равен критическому:  $\Psi_{\text{внеш}} = \Psi_0/2$ ,  $\Psi_{\text{инд}} = -\Psi_0/2$  и  $\Psi_{\text{внут}} = \Psi_{\text{внеш}} + \Psi_{\text{инд}} = 0$ .

При переходе внешнего магнитного потока через значение  $\Psi_0/2$  (т.е. при переходе величины тока через критическое значение  $I_0$ ) в контуре появляется падение напряжения на контакте Джозефсона, т.е. он приобретает некоторое омическое сопротивление, перестает быть чисто сверхпроводящим и в него может проникнуть магнитное поле. Минимально возможная порция — это квант магнитного потока  $\Psi_0$ . Но только половина этой величины создается внешним магнитным полем, другая же половина должна создаваться индукционным током, величина которого равна критическому значению  $I_0$ . Но теперь ток должен не компенсировать внешнее поле, а удвоить его величину, поэтому он скачком меняет свое направление (рис. 43.20, в), так что  $\Psi_{\text{внеш}} = \Psi_0/2$ ,  $\Psi_{\text{инд}} = \Psi_0/2$  и  $\Psi_{\text{внут}} = \Psi_{\text{внеш}} + \Psi_{\text{инд}} = \Psi$ .

При дальнейшем повышении напряженности внешнего магнитного поля ток должен добавлять меньшую порцию потока, и потому его величина уменьшается. Соответственно, контур снова становится сверхпроводящим и магнитный поток сквозь него опять «замораживается». Когда внешний поток достигает величины  $\Psi_{\text{внеш}} = \Psi_0$ , для сохранения потока  $\Psi_{\text{внут}} = \Psi_0$  через контур ток обязан стать равным нулю (рис. 43.20, г). Для компенсации дальнейшего увеличения внешнего поля ток должен начать идти в первоначальном направлении. Когда внешний магнитный поток приблизится к значению  $\Psi_{\text{внеш}} = 3\Psi_0/2$ , ток снова станет близок к критическому, и магнитные потоки будут иметь значения  $\Psi_{\text{внеш}} = 3\Psi_0/2$ ,  $\Psi_{\text{инд}} = -\Psi_0/2$  и  $\Psi_{\text{внут}} = \Psi_{\text{внеш}} + \Psi_{\text{инд}} = \Psi_0$  (рис. 43.20, д). При переходе через критическое значение  $I_0$  внутрь контура проникнет еще один квант магнитного потока ( $\Psi_{\text{внут}} = 2\Psi_0$ , рис. 43.20, е) и ток опять скачком поменяет направление, чтобы создаваемый им поток  $\Psi_{\text{инд}} = \Psi_0/2$  мог добавиться к потоку внешнего поля. При дальнейшем повышении тока через контакт Джозефсона описанные процессы повторяются.

Мы видим, что при внешнем магнитном потоке, состоящем из целого числа квантов потока, состояния сверхпроводников устойчивы, а при полуцелом числе квантов неустойчивы. На рисунке 43.21 показаны графики зависимости магнитного потока через контур и сверхпроводящего тока в нем от потока внешнего магнитного поля. Ступенчатый характер зависимости позволяет наблюдать отдельные кванты магнитного потока, несмотря на их малую величину.

<sup>6</sup> Ю. М. Ципенюк. Физические основы сверхпроводимости. М.: МФТИ, 1996.



**Рис. 43.21.** Зависимость магнитного потока внутри сверхпроводящего контура с контактом Джозефсона (слева) и тока в контуре (справа) от внешнего магнитного потока

### Нестационарный эффект Джозефсона

Если через контакт Джозефсона пропускать внешний постоянный ток  $I > I_0$ , то на контакте возникнет падение напряжения  $U$ . Оно приведет к сдвигу энергетических уровней конденсата  $E_2 - E_1 = -2eU$  (здесь учтено, что куперовская пара имеет двойной электрический заряд). Тогда уравнение (43.45) примет вид

$$\frac{\partial \delta}{\partial t} = \frac{2eU}{\hbar}. \tag{43.46}$$

В этом режиме контакт Джозефсона можно представить себе как параллельное включение обычного омического сопротивления  $R$ , на котором напряжение  $U$  создает нормальный ток  $I_n = U/R$ , и сверхпроводящей ветви, в которой течет ток Джозефсона  $I_s$  (43.43). В рамках этой *резистивной модели* полный ток  $I$  через контакт равен сумме этих двух токов:

$$I = I_0 \sin \delta + \frac{U}{R}. \tag{43.47}$$

Используя (43.46), уравнение (43.47) можем записать в виде дифференциального уравнения для фазы волновой функции:

$$\frac{\partial \delta}{\partial t} = \frac{2eR}{\hbar} (I - I_0 \sin \delta). \tag{43.48}$$

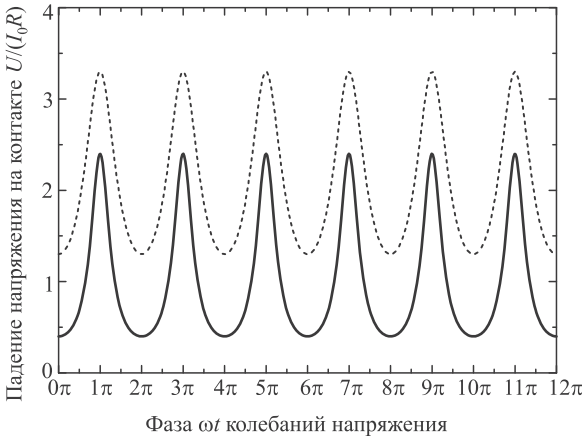
Уравнение (43.48) легко интегрируется после разделения переменных:

$$\int_{\pi/2}^{\delta} \frac{dx}{I - I_0 \sin x} = \frac{2eR}{\hbar} t. \tag{43.49}$$

Здесь мы приняли начальное условие  $\delta(0) = \pi/2$ , т.е. за начало отсчета времени  $t = 0$  взяли момент, когда сверхпроводящий ток достиг максимума:  $I_s(0) = I_0$ . Интеграл в левой части берется, и мы приходим к соотношению

$$\arctg \left( \frac{I \operatorname{tg}(\delta/2) - I_0}{\sqrt{I^2 - I_0^2}} \right) = \arctg \left( \frac{I - I_0}{\sqrt{I^2 - I_0^2}} \right) + \frac{\omega t}{2}, \tag{43.50}$$





**Рис. 43.22.** Колебания напряжения на контакте Джозефсона при пропускании через него постоянного тока  $I = 1, 4I_0$  (сплошная линия) и  $I = 2, 3I_0$  (пунктир)

где введено обозначение

$$\omega = \frac{2eR}{\hbar} \sqrt{I^2 - I_0^2}. \quad (43.51)$$

Приравнивая тангенсы углов в левой и правой частях уравнения (43.50) и применяя формулу для тангенса суммы аргументов  $\text{tg}(\alpha + \beta) = (\text{tg} \alpha + \text{tg} \beta) / (1 - \text{tg} \alpha \cdot \text{tg} \beta)$ , записываем алгебраическое уравнение (43.50) в виде

$$\text{tg}(\delta/2) = \frac{\sqrt{I^2 - I_0^2} + (I - I_0) \text{tg}(\omega t/2)}{\sqrt{I^2 - I_0^2} - (I - I_0) \text{tg}(\omega t/2)}. \quad (43.52)$$

Зная тангенс половинного угла, мы можем найти его синус:

$$\sin \delta = \frac{2 \text{tg}(\delta/2)}{1 + \text{tg}^2(\delta/2)} = \frac{I + I_0 - (I - I_0) \text{tg}^2(\omega t/2)}{I + I_0 + (I - I_0) \text{tg}^2(\omega t/2)}. \quad (43.53)$$

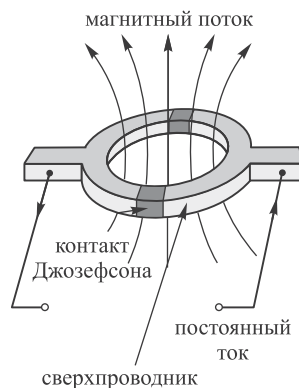
Подставляя (43.53) в (43.47), получаем искомое выражение для падения напряжения на контакте Джозефсона:

$$U = R \frac{I^2 - I_0^2}{I + I_0 \cos \omega t}. \quad (43.54)$$

График зависимости падения напряжения  $U$  от времени показан на рис. 43.22 для двух разных значений тока  $I = 1, 4I_0$  и  $I = 2, 3I_0$ . Видно, что амплитуда колебаний напряжения на контакте растет с ростом силы пропускаемого тока. При  $I \leq I_0$  падение напряжения равно нулю (постоянный ток Джозефсона). Удивительное явление — колебания напряжения при пропускании через контакт Джозефсона постоянного тока — называется **джозефсоновской генерацией**. Впервые это явление наблюдали экспериментально в 1965 г. советские ученые И. К. Янсон, В. М. Свистунов и И. Д. Дмитренко.

Легко найти среднее значение  $\langle U \rangle$  напряжения на контакте, интегрируя (43.54) по периоду:

$$\langle U \rangle = R(I^2 - I_0^2) \frac{\omega}{2\pi} \int_0^{2\pi/\omega} \frac{dt}{I + I_0 \cos \omega t} = R \sqrt{I^2 - I_0^2}. \quad (43.55)$$



**Рис. 43.23.** Схема двухконтактного СКВИДа, работающего на постоянном токе

С учетом определения частоты осцилляций (43.51) этот результат можно представить в виде

$$\langle U \rangle = \frac{\hbar\omega}{2e}, \quad (43.56)$$

допускающем очевидную интерпретацию: коль скоро разница в энергиях куперовских пар по обе стороны от зазора составляет  $2e\langle U \rangle$ , этот излишек энергии может быть сброшен при туннельном переходе куперовской пары через контакт лишь при излучении кванта электромагнитного излучения частотой  $\omega$ .

Свойства джозефсоновского контакта используются в СКВИДах — **сверхпроводящих квантовых интерферометрах** (от англ. *SQUID* — *Superconducting Quantum Interference Device*). СКВИД представляет собой сверхчувствительный магнитометр. Высокочастотный ВЧ-СКВИД работает на переменном токе, его принцип действия основан на нестационарном эффекте Джозефсона и использовании джозефсоновской генерации. Двухконтактный СКВИД работает на постоянном токе и представляет собой контур из сверхпроводника, разорванный в двух местах контактами Джозефсона (рис. 43.23). Если магнитного потока через контур нет, обе ветви сверхпроводящего тока эквивалентны. При наличии магнитного потока через контур ток в одной ветви усиливается, а в другой ослабляется — возникает интерференция джозефсоновских токов, доступная измерению. СКВИДы нашли практическое применение в медицине при анализе работы мозга (магнитоэнцефалография — МЭГ) и сердца (магнитокардиография — МКГ), при изучении земной коры (магниторазведка), в фундаментальных физических исследованиях.

## Контрольные вопросы

1. Перечислите основные особенности явлений сверхтекучести и сверхпроводимости.
2. В чем существенное различие в поведении обычных проводников и сверхпроводников при стремлении температуры к абсолютному нулю?
3. Что такое критическая температура и критическое магнитное поле для сверхпроводников?
4. Что вы знаете о ВТСП? Где могло бы использоваться это явление?

5. Что такое эффект Мейснера? В чем состоит основное отличие сверхпроводника от идеального «классического» проводника?
6. В чем отличие сверхпроводников 1-го и 2-го рода?
7. Что такое вихри Абрикосова?
8. Что представляет собой явление квантования магнитного потока?
9. Что такое электрон-фононное взаимодействие?
10. В чем заключается изотопический эффект и какие следствия вытекают из него для теории сверхпроводимости?
11. Что представляет собой куперовская пара?
12. Поясните в общих чертах роль энергетической щели при возникновении сверхтекучести и сверхпроводимости.
13. В каких других физических системах важную роль играет наличие энергетической щели?
14. Каковы механизмы возникновения туннельного тока при контакте сверхпроводника с нормальным металлом и другим сверхпроводником?
15. Что представляет собой контакт Джозефсона?
16. В чем заключаются стационарный и нестационарный эффекты Джозефсона?
17. Что такое СКВИД?

## Часть VIII

# Основы физики микромира

*Нет вещества — есть круговерти силы;  
Нет твердости — есть натяженья струй;  
Нет атома — есть поле напряженья  
(Вихрь малых «не» вокруг большого «да»);  
Нет плотности, нет веса, нет размера —  
Есть функции различных скоростей.  
Все существует разницей давлений,  
Температур, потенциалов, масс;  
Струи времен текут неравномерно;  
Пространство — лишь разнообразье форм.  
Есть не одна, а много математик;  
Мы существуем в Космосе, где все  
Теряется, ничто не создается. . .*

---

М. Волошин. «Космос»



## Глава 44

# Физика атомного ядра

В 1911 г. Э. Резерфорд открыл, что атом, в сущности, представляет собой пустое пространство. До этого физики, как уже упоминалось, считали атом подобием пудинга с изюмом. Предполагалось, что положительный заряд атома («пудинг») распределен по его объему, а вкрапленные в него электроны («изюм») осциллируют вокруг фиксированных центров. Опыты Резерфорда по рассеянию альфа-частиц на металлической фольге показали, что атом имеет маленькое, но массивное положительно заряженное ядро. Наталкиваясь на него, альфа-частицы (ядра гелия) отклонялись на большой угол, а иногда чуть ли не отскакивали назад. Как позже писал сам Резерфорд, «это все равно, как если бы снаряд из 15-дюймового морского орудия попал в папиросную бумагу и отскочил назад».

Открытием Резерфорда мы уже пользовались при изучении квантовой теории атома водорода, когда описывали электрон в кулоновском поле сил притяжения. В модели пудинга электростатический потенциал был бы, конечно, совершенно другим. Однако свойства атомного ядра, помимо его малых размеров и положительного заряда, нами пока не использовались. Настало время заглянуть внутрь ядра и рассмотреть эти свойства. На этом этапе наших знаний о природе для объяснения всего многообразия мира достаточно трех основных «кирпичиков» — элементарных частиц, из которых построены атом и его ядро. Это протон, нейтрон и электрон. В данной главе мы познакомимся с основными свойствами ядра, его радиоактивными распадами, некоторыми применениями ядерной физики в науке и технике.

### 44.1 Состав атомного ядра

Ядра атомов всех элементов состоят из двух видов элементарных частиц — **протонов** ( $p$ ) и **нейтронов** ( $n$ ). Для этих частиц используют общее название **нуклоны** (от лат. *nucleus* — ядро). Ядра обозначаются символом  ${}^A_ZX$ , где  $X$  — химический символ данного элемента (Na, C, O, Al и т. п.) Слева вверху ставится **массовое число**  $A$ , которое равно суммарному числу протонов и нейтронов (т. е. полному числу нуклонов). Слева внизу стоит **атомный номер**  $Z$  ядра, определяющий порядковый номер химического элемента в периодической таблице

Менделеева. Атомный номер  $Z$  — это количество протонов в ядре, так что заряд ядра есть  $+Ze$ . Число нейтронов в ядре равно  $N = A - Z$ . Строго говоря, левый нижний индекс не нужен, так как атомный номер ядра легко определяется по химическому символу элемента, и мы иногда будем его опускать (например, для урана  $^{235}\text{U}$  по таблице Менделеева находим:  $Z = 92$ ).

До открытия нейтрона (Дж. Чедвик, 1932, Нобелевская премия в 1935 г.) полагали, что ядро состоит из  $A$  протонов и  $(A - Z)$  электронов. Однако в конце 1920-х годов эта теория столкнулась с так называемой азотной катастрофой: ядро азота ( $Z = 7$ ,  $A = 14$ ) должно было бы состоять из 14 протонов и 7 электронов. Всего в составе получалась 21 частица, и у каждой спин был равен  $1/2$ . Следовательно, спин ядра также должен был бы быть полуцелым, а из опытов следовало, что он равен единице. После открытия нейтрона стало понятно, что в ядре азота имеется 7 протонов и 7 нейтронов и эти 14 частиц с полуцелым спином приводят к целому спину ядра.

Рассмотрим основные характеристики нуклонов. Заряд протона  $+e$  положителен и численно равен заряду электрона, масса его  $m_p = 1,6726 \cdot 10^{-27}$  кг. (Для сравнения напомним массу электрона:  $m_e = 9,10939 \cdot 10^{-31}$  кг, т.е.  $m_p = 1836m_e$ .) Обычно при использовании связи энергии и массы  $E = mc^2$  массы элементарных частиц указывают в энергетических единицах ( $\text{МэВ}/c^2$ ); тогда  $m_p = 938,27 \text{ МэВ}/c^2$ ,  $m_e = 0,511 \text{ МэВ}/c^2$ . Спин протона равен половине ( $s = 1/2$ ), а его собственный магнитный момент  $\mu_p = +2,79\mu_N$ , где **ядерный магнетон**  $\mu_N = e\hbar/(2m_p) = 5,05 \cdot 10^{-27}$  Дж/Тл  $= 3,152 \cdot 10^{-14}$  МэВ/Тл. Из-за массивности протона ядерный магнетон в 1836 раз меньше магнетона Бора. Размеры протона составляют  $r_p \sim 1,3 \cdot 10^{-15}$  м.

Нейтрон имеет нулевой электрический заряд и массу  $m_n = 1,6749 \cdot 10^{-27}$  кг  $= 939,56 \text{ МэВ}/c^2$ . Отметим, что нейтрон слегка тяжелее протона, их разность масс  $\delta m = m_n - m_p = 1,3 \text{ МэВ}/c^2$ , что составляет примерно две с половиной массы электрона. Спин нейтрона также равен половине ( $s = 1/2$ ), а собственный магнитный момент (несмотря на отсутствие электрического заряда) отличен от нуля, но имеет отрицательный знак и равен  $\mu_n = -1,91\mu_N \approx -2\mu_p/3$ . Существование магнитного момента у нейтрона и аномальная величина магнитного момента протона ( $2,79\mu_N$  вместо одного ядерного магнетона) связаны с неточностью нуклонов, о которой мы поговорим позже. В свободном состоянии нейтрон неустойчив (нестабилен): он самопроизвольно распадается на протон ( $p$ ), электрон ( $e^-$ ) и **электронное антинейтрино** (символ  $\bar{\nu}_e$ ):  $n \rightarrow p + e^- + \bar{\nu}_e$ . Электронное антинейтрино — элементарная частица с ничтожно малой массой (меньше  $0,3 \text{ эВ}/c^2$ ) и нулевым зарядом. Среднее время жизни нейтрона равно  $\tau_n = 14,8$  мин. Так как энергия покоя нейтрона  $E_{0,n} = m_n c^2 = 939,56 \text{ МэВ}$  превышает сумму энергий покоя ( $E_{0,p} + E_{0,e} = 938,27 \text{ МэВ} + 0,511 \text{ МэВ} = 938,78 \text{ МэВ}$ ) протона и электрона, образующихся при распаде нейтрона, на величину  $\Delta E = E_{0,n} - E_{0,p} - E_{0,e} = 0,78 \text{ МэВ}$ , этот избыток энергии  $\Delta E$  выделяется в виде кинетической энергии образовавшихся частиц.

Ядра с одинаковым числом протонов  $Z$ , но с разными массовыми числами  $A$  называют **изотопами**. Имея одинаковый заряд, они принадлежат тому же химическому элементу. Например, водород имеет три изотопа:

${}^1_1\text{H}$  — обычный водород ( $Z = 1, A = 1, N = 0$ );

${}^2_1\text{H}$  — тяжелый водород (дейтерий) ( $Z = 1, A = 2, N = 1$ );

${}^3_1\text{H}$  — сверхтяжелый водород (тритий) ( $Z = 1, A = 3, N = 2$ ).

Обычный водород и дейтерий стабильны, а тритий радиоактивен (самопроизвольно распадается). Большинство химических элементов имеет несколько стабильных изотопов. Наличие различных изотопов у химических элементов обуславливает нецелую величину атомной массы в периодической таблице Менделеева, поскольку в естественном состоянии элемент представляет собой смесь различных стабильных изотопов.

Ядра с одинаковым массовым числом  $A$  называются *изобарами* (например,  ${}^{40}_{18}\text{Ar}$  и  ${}^{40}_{20}\text{Ca}$ ). Ядра, имеющие одинаковое число нейтронов  $N = A - Z$ , носят название *изотонов* ( ${}^{13}_6\text{C}$  и  ${}^{14}_7\text{N}$ ).

## 44.2 Физическая природа ядерных сил

Устойчивость ядер (по крайней мере тех, что встречаются в природе в естественном состоянии, а не «изготовлены» физиками) свидетельствует об интенсивном взаимодействии нуклонов, которое носит характер притяжения. Притяжение между нуклонами удерживает их на малых расстояниях  $\sim 10^{-15}$  м, несмотря на кулоновское отталкивание заряженных протонов. Это взаимодействие нуклонов называется *сильным ядерным взаимодействием*.

### Свойства сильного ядерного взаимодействия.

- Ядерные силы являются короткодействующими, их радиус действия имеет порядок  $10^{-15}$  м, причем на расстояниях, существенно меньших  $10^{-15}$  м, притяжение нуклонов сменяется сильным отталкиванием.
- Ядерные силы зарядовнезависимы: взаимодействия пар  $p - p$ ,  $n - n$ ,  $n - p$  имеют одинаковую интенсивность.
- Ядерные силы не являются центральными: эти силы не направлены вдоль прямой, соединяющей центры взаимодействующих нуклонов.
- Ядерные силы зависят от взаимной ориентации спинов нуклонов. Например, ядро тяжелого водорода (дейтрон) образуется только в том случае, если нейтрон и протон имеют спины, параллельные друг другу (т. е. спин дейтрона равен единице).
- Ядерные силы имеют свойство насыщения: каждый нуклон в ядре взаимодействует не со всеми нуклонами, а лишь с ближайшими соседями. Как следствие, энергия связи одного нуклона с ядром примерно постоянна для всех ядер с большим числом нуклонов.
- Плотность ядерного вещества является постоянной: объем ядра пропорционален числу образующих его нуклонов. Это также является следствием свойства насыщения ядерных сил: при увеличении числа нуклонов ядро не сжимается все сильнее и сильнее, как это происходило бы с далекодействующими силами притяжения, подобными силам гравитации.



Согласно *квантовой электродинамике*, о которой мы поговорим в следующей главе, электромагнитные силы возникают, когда заряженные частицы обмениваются фотонами. Поскольку масса фотонов равна нулю, радиус действия электромагнитных сил бесконечен. В 1935 г. японский физик Х. Юкава предположил, что переносчиком ядерного взаимодействия являются гипотетические «тяжелые фотоны».

Их массу  $m$  можно оценить на основе соотношения неопределенностей, зная радиус действия ядерных сил  $r_{\text{яд}} \sim 1,3 \cdot 10^{-15}$  м. Пусть покоящийся нуклон испустил «тяжелый фотон» массой  $m$ . Произошло изменение энергии системы  $\Delta E \sim mc^2$ , но сам нуклон ничуть не изменился. Согласно принципу неопределенности, такое нарушение закона сохранения энергии принципиально не будет наблюдаться в течение времени  $\Delta t \sim \hbar/\Delta E = \hbar/(mc^2)$ , после чего испущенная частица должна быть поглощена другим нуклоном. Это и будет означать, что наши два нуклона провзаимодействовали друг с другом. Но за время  $\Delta t$  «тяжелый фотон» может преодолеть расстояние, не большее

$$r_{\text{яд}} = c\Delta t \sim \frac{\hbar}{mc} = \lambda_{m,C}. \quad (44.1)$$

Таким образом, комптоновская длина волны переносчика сильного ядерного взаимодействия по порядку величины равна радиусу действия порождаемых им сил. Записывая это соотношение в виде  $mc^2 \sim \hbar c/r_{\text{яд}}$  и учитывая значение комбинации фундаментальных постоянных  $\hbar c$  в энергетических единицах ( $\hbar c = 193,327 \cdot 10^{-15}$  МэВ · м), оцениваем массу переносчиков сильных ядерных взаимодействий:  $m \sim 193,327/1,3 = 150$  МэВ/ $c^2$ . Носители ядерных сил (так называемые  $\pi$ -мезоны) были открыты в 1947 г. в экспериментах Дж. Оккиалини и С. Пауэлла (последний был удостоен в 1950 г. Нобелевской премии «за разработку фотографического метода исследования ядерных процессов и открытие мезонов, осуществленное с помощью этого метода»). Оказалось, что существуют заряженные  $\pi$ -мезоны ( $\pi^+$  и  $\pi^-$  с зарядами  $\pm e$  и массами  $m_{\pi^\pm} = 139,57$  МэВ/ $c^2$ ) и нейтральный  $\pi^0$ -мезон с массой  $m_{\pi^0} = 134,97$  МэВ/ $c^2$ . Спин всех  $\pi$ -мезонов (заряженных и нейтрального) равен нулю ( $s = 0$ ). Все три  $\pi$ -мезона являются нестабильными. Времена жизни составляют:  $\tau_{\pi^\pm} = 2,6 \cdot 10^{-8}$  с,  $\tau_{\pi^0} = 0,8 \cdot 10^{-16}$  с. Название «тяжелые фотоны» более не используется (так, скорее уж, можно было бы назвать переносчиков слабого ядерного взаимодействия — массивные  $W^\pm$  и  $Z^0$  бозоны, о которых речь пойдет в гл. 45). Ну а их «изобретатель» Х. Юкава стал в 1949 г. лауреатом Нобелевской премии «за предсказание существования мезонов на основе теоретической работы по ядерным силам».

Процессы испускания частицей квантов поля, когда сами кванты принципиально не могут наблюдаться, но либо поглощаются через короткое время этой же частицей, либо какой-нибудь другой, взаимодействующей с первой, называются *виртуальными*. Такие процессы лежат в основе всех теорий взаимодействия фундаментальных частиц, иного механизма современная физика не знает.

Заметим, что при испускании заряженных  $\pi^\pm$ -мезонов заряд частицы меняется, т. е. протоны и нейтроны взаимно превращаются друг в друга. Этого не происходит при обмене нуклонов нейтральным  $\pi^0$ -мезоном.

Таким образом, в результате виртуальных процессов

$$p \rightleftharpoons n + \pi^+, \quad p \rightleftharpoons p + \pi^0, \quad n \rightleftharpoons p + \pi^-, \quad n \rightleftharpoons n + \pi^0$$

нуклоны оказываются окруженными облаком виртуальных  $\pi$ -мезонов, образующих поле ядерных сил. Облако  $\pi$ -мезонов иногда называют «мезонной шубой». Испускание и поглощение виртуальных  $\pi$ -мезонов нуклонами приводит к **сильному ядерному взаимодействию** между ними, которое осуществляется по следующей схеме: при взаимодействии, скажем, протона и нейтрона последний может испустить  $\pi^-$ -мезон, превратившись в протон. Начальный же протон, поглощая этот  $\pi^-$ -мезон, превращается в нейтрон. В конечном состоянии остаются те же самые частицы:

$$p + n \rightarrow p + (\pi^- + p) \rightarrow (p + \pi^-) + p \rightarrow n + p.$$

Нетрудно записать и другие подобные виртуальные реакции взаимопревращения нуклонов с участием  $\pi^+$ - и  $\pi^0$ -мезонов.

На вопрос, почему обмен  $\pi$ -мезонами приводит к притяжению нуклонов, а обмен фотоном — к отталкиванию одноименно заряженных электронов, здесь ответить невозможно. Отметим лишь, что это связано со спином квантов переносчиков взаимодействия:  $s = 0$  для  $\pi$ -мезонов и  $s = 1$  для фотонов.

Наличие «мезонной шубы» позволяет объяснить аномальную величину магнитного момента протона ( $\mu_p = 2,79\mu_N$ ) и ненулевое значение магнитного момента нейтрона ( $\mu_n = -1,91\mu_N$ , где минус указывает на то, что направления собственных механического (спинового) и магнитного моментов противоположны). Нейтрон вследствие реакции  $n \rightleftharpoons p + \pi^-$  часть времени проводит в виртуальном состоянии «протон +  $\pi^-$ -мезон», причем орбитальное движение  $\pi^-$  приводит к возникновению у нейтрона отрицательного магнитного момента. Орбитальное движение  $\pi^+$ -мезона в виртуальном состоянии протона ( $p \rightleftharpoons n + \pi^+$ ) объясняет аномальный магнитный момент протона.

## 44.3 Масса и дефект массы ядра

Можно ли подсчитать массу ядра по его составу? На первый взгляд нет ничего проще. Возьмем в качестве примера дейтрон — ядро тяжелого водорода  ${}^2_1\text{H}$ , состоящее из протона и нейтрона. Суммарная масса этих частиц равна  $m_p + m_n = = 938,27 + 939,56 = 1877,83 \text{ МэВ}/c^2$ . Эксперимент же дает для массы дейтрона значение  $m_d = 1875,63 \text{ МэВ}/c^2$ , т. е. на  $2,22 \text{ МэВ}/c^2$  меньше. Эта разность, называемая **дефектом массы**, — проявление знаменитой формулы  $E = mc^2$ . Действительно, при объединении протона и нейтрона в дейтрон выделилась какая-то энергия  $E_{\text{св}}$ . Но выделение энергии эквивалентно уменьшению массы системы. Поэтому масса дейтрона (и любого другого ядра) меньше массы его компонентов на величину  $E_{\text{св}}/c^2$ . Чтобы снова разделить дейтрон на свободные протон и нейтрон, точно такую же энергию надо затратить. Это и есть **энергия связи** дейтрона  $E_{\text{св}} = 2,22 \text{ МэВ}$ .

В общем случае, при объединении нуклонов в ядро, масса  $M$  ядра всегда меньше (из-за дефекта массы) суммы масс  $Zm_p + (A - Z)m_n$  входящих в него

частиц на величину

$$\Delta_M = Zm_p + (A - Z)m_n - M. \quad (44.2)$$

Соответственно энергия покоящегося ядра  $E = Mc^2$  меньше суммарной энергии невзаимодействующих покоящихся нуклонов на величину энергии связи

$$E_{\text{св}} = \Delta_M c^2 = [Zm_p + (A - Z)m_n - M]c^2. \quad (44.3)$$

♦ **Энергия связи** определяет величину работы, которую необходимо совершить, чтобы удалить нуклоны друг от друга на расстояния, превышающие радиус действия ядерных сил.

Энергия связи, приходящаяся на один нуклон ядра ( $E_{\text{св}}/A$ ), называется **удельной энергией связи** нуклонов в ядре. ♦

В дейтроне она равна примерно 1,1 МэВ/нуклон.

В ядерной физике часто удобнее пользоваться не массами «голых» ядер, а массами нейтральных атомов, включая и массу всех электронов (дефект массы из-за связи электронов с ядром ничтожен, и им пренебрегают). Естественно, формула (44.3) для энергии связи останется прежней, только теперь  $M$  — это масса атома, а вместо массы протона  $m_p$  надо использовать массу нейтрального атома водорода  ${}^1\text{H}$ . Кроме того, массы часто выражают в так называемых **атомных единицах массы** — **а. е. м.** По определению 1 а. е. м. =  $1/12$  массы атома углерода  ${}^{12}\text{C}$ . Установим связь с обычными единицами массы. Один моль атомов  ${}^{12}\text{C}$  имеет массу 12 г, следовательно, 1 а. е. м. =  $10^{-3}/N_A = 1,66054 \cdot 10^{-27}$  кг. Учитывая, что энергия  $E$ , соответствующая массе  $m = 1$  кг, равна  $E = mc^2 = 8,98755 \cdot 10^{16}$  Дж =  $5,60959 \cdot 10^{35}$  эВ, находим, что 1 а. е. м. =  $931,495$  МэВ/ $c^2$ . В этих единицах  $m_p = 1,007276$  а. е. м.,  $m_n = 1,008665$  а. е. м., а масса нейтрального атома водорода  $m_{\text{H}} = 1,007825$  а. е. м.

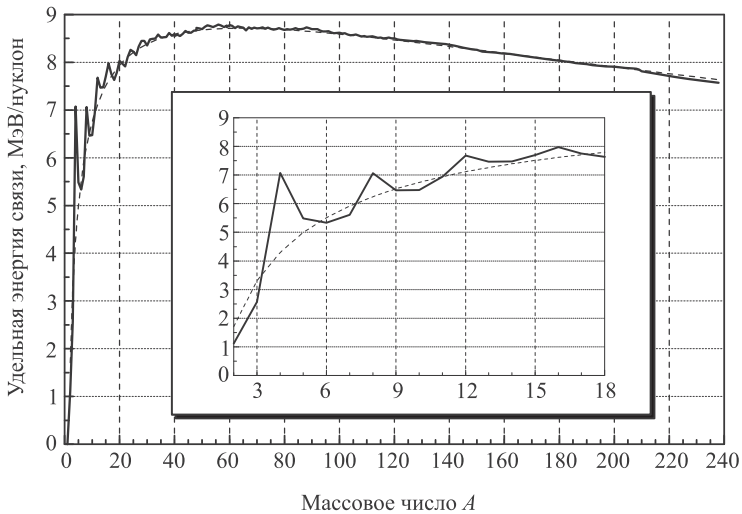
**Задача 44.65.** Вычислить дефект массы, энергию связи и удельную энергию связи нуклонов в ядре гелия  ${}^4_2\text{He}$ . Масса атома гелия составляет 4,00260 а. е. м.

**Решение.** В состав ядра гелия входят два протона ( $Z = 2$ ) и два нейтрона ( $N = A - Z = 2$ ). Отсюда сначала находим дефект массы:  $\Delta_M = 2 \times 1,00783 + 2 \times 1,00867 - 4,00260 = 0,0304$  а. е. м. Для энергии связи теперь получаем  $E_{\text{св}} = \Delta_M c^2 = 0,0304 \times 931,495 = 28,32$  МэВ. Удельная энергии связи при четырех нуклонах в ядре равна  $28,32/4 = 7,08$  МэВ/нуклон. Для сравнения: мы видели, что энергия связи валентных электронов в атомах имеет величину порядка десятка электронвольт, т. е. она почти в миллион раз меньше удельной энергии связи нуклонов в ядрах. ■

Заметим, что выражения (44.2) для дефекта массы и (44.3) для энергии связи можно использовать не только по отношению к расщеплению ядер на составляющие, но и для других ядерных реакций и распадов. Соответствующие модификации легко понять: если ядро массой  $M$  разваливается на ядра меньших масс  $M_1$  и  $M_2$ , то дефект массы и энергию связи следует определять по формулам

$$\Delta_M = M_1 + M_2 - M, \quad E_{\text{св}} = \Delta_M c^2 = (M_1 + M_2 - M)c^2. \quad (44.4)$$

Для иллюстрации применения этих формул рассмотрим поучительный пример последовательного расщепления ядра  ${}^4\text{He}$ , позволяющий понять, из чего складываются 28,3 МэВ, которые мы нашли для энергии связи ядра гелия. На первом этапе оторвем от ядра протон:  ${}^4\text{He} \Rightarrow {}^3\text{H} + {}^1\text{H}$ . Масса изотопа



**Рис. 44.1.** Удельная энергия связи нуклонов в ядрах: рассчитанная по формуле (44.21) штриховая кривая практически сливается с экспериментальной сплошной; различие видно в области пиков для легких элементов в интервале от  $A = 2$  до  $A = 18$  (показано на врезке)

${}^3\text{He}$  равна 3,0161 а.е.м. Поэтому дефект массы в этом случае составляет  $\Delta_{M_1} = 3,0161 + 1,0078 - 4,0026 = 0,0213$  а.е.м., а энергия отрыва протона  $E_1 = 0,0213 \times 931,5 = 19,84$  МэВ. На втором этапе оторвем от трития  ${}^3\text{H}$  нейтрон:  ${}^3\text{H} \Rightarrow {}^2\text{H} + n$ . Масса атома дейтерия равна 2,0141 а.е.м., так что дефект массы  $\Delta_{M_2} = 2,0141 + 1,0087 - 3,0161 = 0,0067$  а.е.м., а энергия отрыва второго нуклона  $E_2 = 0,0067 \times 931,5 = 6,24$  МэВ. Наконец, оторвем еще один нейтрон от дейтрона:  ${}^2\text{H} \Rightarrow {}^1\text{H} + n$ . Дефект массы  $\Delta_{M_3} = 1,0078 + 1,0087 - 2,0141 = 0,0024$  а.е.м., энергия отрыва  $E_3 = 0,0024 \times 931,5 = 2,23$  МэВ. Складывая найденные величины, получаем в точности энергию связи ядра гелия:  $E_1 + E_2 + E_3 = 19,84 + 6,24 + 2,23 = 28,3$  МэВ. Примерно так же будет выглядеть расчет при другом порядке «расщепления» ядра гелия: трудно оторвать первую частицу, легче — последующие.

Очень информативным является график зависимости удельной энергии связи  $E_{\text{св}}$  от массового числа  $A$  (рис. 44.1). Наибольшей высоты кривая достигает на ядре  ${}^{62}\text{Ni}$ , близко к вершине лежат ядра железа  ${}^{56}\text{Fe}$  и соседних металлов. Энергия связи этих ядер составляет 8,7 МэВ/нуклон. Примерное постоянство энергии связи в средней части графика 44.1 говорит о насыщении ядерных сил: не все нуклоны взаимодействуют друг с другом. В противном случае энергия связи была бы пропорциональна числу пар ( $\sim A^2$ ), а удельная энергия связи тогда бы росла:  $E_{\text{св}}/A \sim A$ . Насыщение очевидным образом связано с короткодействием ядерных сил. При переходе к еще более тяжелым ядрам с большими  $A$  удельная энергия связи постепенно уменьшается, и для урана она составляет 7,5 МэВ/нуклон. Уменьшение устойчивости тяжелых ядер объясняется тем, что кулоновские силы отталкивания между протонами имеют бесконечный радиус действия, что при большом числе протонов приводит к развалу ядра.

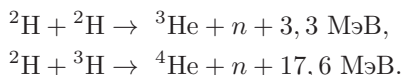
Указанная зависимость  $E_{\text{св}}/A$  от  $A$  дает принципиальную возможность получать энергию при превращениях атомных ядер. В самом деле, если надо затратить энергию  $E_{\text{св}}$  для того, чтобы развалить ядро на составляющие нуклоны, то в процессе собирания этих нуклонов в ядро такое же количество энергии будет выделяться. Полезный выход энергии получится тогда, когда в процессе превращения выделяемая энергия превысит затраченную, т. е. когда среди конечных продуктов реакции окажутся ядра с большей энергией связи, нежели у начальных ядер. Иными словами, мы должны в результате реакции подниматься по кривой, показанной на рис. 44.1. Такой подъем возможен, если стартовать с тяжелых ядер на правом краю кривой и переходить к ядрам средней массы в центре кривой. Или же надо начинать с совсем легких ядер на левом краю кривой и двигаться в сторону более тяжелых. Физически первый процесс означает **деление тяжелых ядер** на несколько более легких, а во втором происходит **слияние (синтез) легких ядер** в одно ядро. Оба процесса сопровождаются выделением большого количества энергии.

Например, одна из возможных *реакций деления* происходит в ядерном реакторе, когда нейтрон попадает в ядро изотопа урана  ${}_{92}^{235}\text{U}$  и разваливает его на несколько осколков:



где  $Q_f$  — выделяющаяся энергия. Используя рис. 44.1, оценим энерговыделение в реакциях деления урана. Поскольку удельная энергия связи урана составляет 7,5 МэВ/нуклон, а удельная энергия связи осколков в середине рисунка примерно равна 8,7 МэВ/нуклон, выделяемая энергия составит около 1 МэВ/нуклон. Умножая на число нуклонов в ядре урана, получаем, что полное энерговыделение примерно равно 240 МэВ.

В земных условиях удалось осуществить две реакции синтеза — слияние ядер тяжелых изотопов водорода в ядро гелия:



В первой реакции выделяется энергия  $3,3/4 = 0,8$  МэВ/нуклон, во второй —  $17,6/5 = 3,5$  МэВ/нуклон. К сожалению, обе реакции пока осуществлены лишь в военных целях — при термоядерных взрывах. Для сравнения укажем, что при лежащей в основе нашей цивилизации реакции горения



выделяется энергия  $4,2/(12+32) = 0,095$  эВ/нуклон. При пересчете в обычные единицы это и даст теплотворную способность топлива (например, теплота сгорания бензина равна 44 МДж/кг). «Теплотворная» же способность ядерного топлива, как мы видели, намного выше, что и предопределяет развитие ядерной энергетики. Различие в миллионы раз связано с тем, что при химических реакциях пространственная область, в которой происходит перегруппировка частиц, определяется размерами молекул, а в случае ядерных превращений — гораздо меньшими размерами ядра.

## 44.4 Модели атомного ядра

В этом разделе мы очень коротко познакомимся с двумя моделями ядра, применяемыми для описания хода кривой энергии связи ядер.

### Модель жидкой капли

Для начала проведем оценку плотности атома водорода. Масса протона  $m_p = 1,67 \cdot 10^{-27}$  кг, радиус атома  $a_B = 0,5 \cdot 10^{-10}$  м. Тогда средняя плотность атома равна  $\rho = \frac{m_p}{4\pi a_B^3/3} \approx 3,2 \cdot 10^3$  кг/м<sup>3</sup> = 3,2 кг/дм<sup>3</sup>. Плотности обычных

для нас веществ имеют близкие значения: жидкости — 0,7–1,8 кг/дм<sup>3</sup>, уголь — 1,4 кг/дм<sup>3</sup>, алюминий — 2,7 кг/дм<sup>3</sup>, гранит — 2,8 кг/дм<sup>3</sup>, железо — 7,8 кг/дм<sup>3</sup>. Вывод: атомы в веществе плотно упакованы, так что плотность вещества примерно равна плотности атомов.

Вполне естественно предположить, что нуклоны в ядрах тоже плотно упакованы, и даже намного плотнее, чем атомы, поскольку ядра представляют собой гораздо более однородную структуру, нежели атомы: разные вещества построены из различных атомов, в то время как все ядра состоят из одинаковых нуклонов. Оценим среднюю плотность протона, исходя из известного нам радиуса действия ядерных сил (или, что то же самое, радиуса протона  $r_p = 1,3 \cdot 10^{-15}$  м):

$\rho_p = \frac{m_p}{4\pi r_p^3/3} = 1,8 \cdot 10^{17}$  кг/дм<sup>3</sup>. Примерно такое же значение должна иметь

плотность ядерной материи во всех ядрах. Это и есть первое предположение, лежащее в основе модели жидкой капли, развитой в 1930-е гг. (Г. А. Гамов, Н. Бор, Я. И. Френкель). Плотность ядерной материи пропорциональна числу нуклонов  $A$  и обратно пропорциональна объему ядра, т. е. величине  $r_\text{я}^3$ , где  $r_\text{я}$  — радиус ядра. Коль скоро отношение  $A/r_\text{я}^3 = \text{const}$ ,  $r_\text{я} \sim A^{1/3}$ . Коэффициент пропорциональности должен быть равен радиусу протона, чтобы получилось правильное значение при  $A = 1$ . Таким образом, в первом приближении ядро можно рассматривать как шар радиусом

$$\boxed{r_\text{я} = r_p A^{1/3}}, \quad (44.5)$$

где  $r_p = 1,3 \cdot 10^{-15}$  м — радиус протона.

**Задача 44.66.** Нейтронная звезда — это объект, чья плотность примерно равна плотности ядерной материи. Вычислить, какой радиус имело бы Солнце, превратившись в нейтронную звезду. Считать, что при таком превращении масса Солнца  $M_\odot = 1,99 \cdot 10^{30}$  кг не изменилась.

**Решение.** Плотность ядерной материи найдем из данных о протоне:  $\rho_p = \frac{m_p}{4\pi r_p^3/3}$ . Ту же плотность будет иметь Солнце после превращения в нейтронную звезду:  $\rho_p = \frac{M_\odot}{4\pi R_{n*}^3/3}$ . Разделив первое равенство на второе, получим  $1 = m_p R_{n*}^3 / (M_\odot r_p^3)$ , откуда

$$R_{n*} = r_p \sqrt[3]{\frac{M_\odot}{m_p}} = 1,3 \cdot 10^{-15} \sqrt[3]{\frac{1,99 \cdot 10^{30}}{1,67 \cdot 10^{-27}}} = 13,8 \text{ км.} \quad \blacksquare \quad (44.6)$$

### Формула Бете—Вейцекера

Попробуем вывести полуэмпирическую формулу для массы  $M(Z, A)$  ядра. Главным слагаемым в ней, естественно, должна быть полная масса входящих в ядро нуклонов (речь идет не о собственно ядрах, а о массах соответствующих нейтральных атомов):

$$\begin{aligned} M(Z, A) &= 1,007825Z + 1,008665(A - Z) + \dots = \\ &= 1,008665A - 0,000840Z + \dots \end{aligned} \quad (44.7)$$

Здесь массы выражены в атомных единицах массы, а точками показаны члены, описывающие дефект массы.

Второе предположение капельной модели ядра использует короткодействие ядерных сил. Предполагается, что в выражение для массы ядра входит член  $M_{NN} \sim A$ , описывающий взаимодействие нуклонов и пропорциональный массовому числу  $A$ . Это слагаемое должно войти с отрицательным знаком — взаимодействие нуклонов понижает массу ядра, придавая ему стабильность (дефект массы). Если бы ядерные силы были дальнедействующими, то каждый нуклон взаимодействовал бы с каждым, так что это слагаемое в полной массе должно было бы меняться как  $A^2$ . Итак, из капельной модели ядра на данном этапе нашего вывода следует выражение вида

$$M(Z, A) = 1,008665A - 0,000840Z - a_1A + \dots \quad (44.8)$$

Коэффициент пропорциональности  $a_1$ , равно как и коэффициенты при следующих слагаемых, будут определены впоследствии путем сравнения полученной формулы с экспериментальными данными.

Только что введенная поправка приписывает всем нуклонам одинаковую энергию взаимодействия. Но это заведомо не так: нуклоны на поверхности ядра испытывают притяжение только с одной стороны — внутрь капли. Это явление сродни поверхностному натяжению в капле обычной жидкости. Но поверхностное натяжение пропорционально площади поверхности ядра, т. е. мы должны ввести поправку  $M_S \sim r_{\text{я}}^2 \sim A^{2/3}$  в соответствии с формулой (44.5). Это слагаемое должно ослабить действие первой поправки на взаимодействие нуклонов, т. е. оно войдет в выражение для массы с положительным знаком:

$$M(Z, A) = 1,008665A - 0,000840Z - a_1A + a_2A^{2/3} + \dots \quad (44.9)$$

Далее надо отразить в формуле для массы экспериментальный факт, что для легких ядер числа нейтронов и протонов примерно одинаковы:  $Z \approx N = A - Z \approx A/2$ . Проще всего получить это слагаемое, если представить себе ядро как ферми-жидкость протонов и нейтронов. Энергия Ферми протонов  $\varepsilon_{F,p}$  пропорциональна их плотности в степени  $2/3$ :  $\varepsilon_{F,p} \sim (Z/V)^{2/3}$ . Так как объем ядра пропорционален массовому числу  $A$ , получаем, что  $\varepsilon_{F,p} \sim (Z/A)^{2/3}$ . Полная же энергия протонов равна  $E_p = 3Z\varepsilon_{F,p}/5$ , т. е. может быть записана как  $E_p = c_1 Z^{5/3}/A^{2/3}$ . Точное значение коэффициента  $c_1$  нам сейчас не так важно. Совершенно аналогичный вид имеет полная энергия нейтронов, только число протонов  $Z$  надо заменить на число нейтронов  $A - Z$ :  $E_n = c_1 (A - Z)^{5/3}/A^{2/3}$ . Из суммы этих энергий надо вычесть значение суммы при равенстве чисел нейтронов и протонов (т. е. ее значение при  $Z = A/2$ ). Получаем тогда поправку к энергии в виде

$$E_F = \frac{c_1}{A^{2/3}} \left[ Z^{5/3} + (A - Z)^{5/3} - 2 \left( \frac{A}{2} \right)^{5/3} \right]. \quad (44.10)$$

Введем величину  $\Delta$ , описывающую отклонение числа протонов от числа нейтронов:  $\Delta = (N - Z)/2 = A/2 - Z$ . Подставляя  $Z = A/2 - \Delta$  в (44.10), получаем:

$$\begin{aligned} E_F &= \frac{c_1}{A^{2/3}} \left[ \left( \frac{A}{2} - \Delta \right)^{5/3} + \left( \frac{A}{2} + \Delta \right)^{5/3} - 2 \left( \frac{A}{2} \right)^{5/3} \right] = \\ &= \frac{c_1 A}{2^{5/3}} \left[ \left( 1 - \frac{2\Delta}{A} \right)^{5/3} + \left( 1 + \frac{2\Delta}{A} \right)^{5/3} - 2 \right]. \end{aligned} \quad (44.11)$$

Поскольку  $\Delta$  невелико, мы разложим выражение в квадратных скобках в ряд Тейлора по степеням  $\Delta$ . Первые два члена разложения взаимно компенсируются, и мы оставим в разложении только первый неисчезающий член, пропорциональный  $(\Delta/A)^2$ . В результате получим, что  $E_F \sim A(\Delta/A)^2 = \Delta^2/A = (A/2 - Z)^2/A$ . Такой же вид будет иметь соответствующая поправка к массе, причем она должна войти с положительным знаком, так как отклонение  $Z$  от  $N = A - Z$  уменьшает стабильность ядра, т. е. повышает его массу:

$$M(Z, A) = 1,008665A - 0,000840Z - a_1A + a_2A^{2/3} + a_3 \frac{(A/2 - Z)^2}{A} + \dots \quad (44.12)$$

Помимо слагаемых, учитывающих различные детали ядерного взаимодействия нуклонов, в выражение для массы должен войти также член, описывающий кулоновское отталкивание протонов. Если представить себе ядро как однородно заряженный по объему шар с зарядом  $Ze$  и радиусом  $r_j$ , то его электростатическая энергия будет равна:

$$E_C = \frac{3}{5} \cdot \frac{(Ze)^2}{4\pi\epsilon_0 r_j} = \frac{3}{5} \cdot \frac{e^2}{4\pi\epsilon_0 r_p} \cdot \frac{Z^2}{A^{1/3}} = (0,6646 \text{ МэВ}) \cdot \frac{Z^2}{A^{1/3}}. \quad (44.13)$$

Переводя коэффициент из энергетических единиц в а. е. м. ( $0,6646/931,5 = 0,000713$ ), получаем поправку на кулоновское взаимодействие протонов в ядре. Оно расталкивает протоны, т. е. уменьшает стабильность ядра и приводит к повышению его массы:

$$\begin{aligned} M(Z, A) &= 1,008665A - 0,000840Z - a_1A + a_2A^{2/3} + \\ &+ a_3 \frac{(A/2 - Z)^2}{A} + 0,000713 \frac{Z^2}{A^{1/3}}. \end{aligned} \quad (44.14)$$

Коэффициенты  $a_i$  в (44.14) определяются при сравнении этой формулы со всем массивом экспериментальных данных. Чтобы остаться в рамках алгебры и не прибегать к программам компьютерной обработки экспериментальных данных, мы для демонстрации эффективности модели используем свойства только нескольких ядер, которые выбраны достаточно случайным образом.

Сначала применим (44.14) для определения связи  $Z$  и  $A$  в наиболее стабильных ядрах. Для этого надо найти минимум  $M(Z, A)$  по переменной  $Z$ , т. е. решить уравнение  $\partial M(Z, A)/\partial Z = 0$ . Итак, дифференцируем  $M(Z, A)$ :

$$\frac{\partial M(Z, A)}{\partial Z} = -0,000840 - 2a_3 \frac{(A/2 - Z)}{A} + 2 \cdot 0,000713 \frac{Z}{A^{1/3}} = 0. \quad (44.15)$$

Подставим сюда значения  $Z, A$  для наиболее распространенного изотопа золота  $^{197}_{79}\text{Au}$ . Для определения  $a_3$  имеем тогда уравнение

$$-0,000840 - 0,19797a_3 + 0,01936 = 0, \quad (44.16)$$



откуда следует значение  $a_3 = 0,0936$  а. е. м. Подставляя это значение в (44.15), находим выражение

$$Z = \frac{A}{1,98 + 0,015A^{2/3}}, \quad (44.17)$$

дающее связь между числами протонов и нейтронов в наиболее стабильных ядрах. Для легких ядер, при малых  $A$ , действительно получаем  $Z \approx A/2$ , для тяжелых же ядер число нейтронов начинает превышать число протонов. Читатель, вооружившись таблицей Менделеева, может поиграть с этой формулой и убедиться, что она достаточно хорошо описывает свойства ядер.

Итак, нам осталось найти два коэффициента в выражении для массы ядер. Снова используем данные по золоту:  $M(79,197) = 196,966543$  а. е. м. Подставляя данные в (44.14), получаем одно уравнение:

$$197a_1 - 33,9a_2 = 2,62. \quad (44.18)$$

Для получения второго уравнения возьмем элемент из начала таблицы Менделеева — литий  ${}^7_3\text{Li}$ :  $M(3,7) = 7,016003$  а. е. м. Снова используем (44.14) и приходим к уравнению

$$7a_1 - 3,66a_2 = 0,0488. \quad (44.19)$$

Совместное решение уравнений (44.18), (44.19) дает значения  $a_1 = 0,0164$ ,  $a_2 = 0,0180$ . Таким образом, наша формула для массы ядер приобретает вид

$$M(Z, A) = 1,008665A - 0,000840Z - 0,0164A + 0,0180A^{2/3} + 0,0936 \frac{(A/2 - Z)^2}{A} + 0,000713 \frac{Z^2}{A^{1/3}}. \quad (44.20)$$

Вычитая из  $M(Z, A)$  массы нуклонов, получаем выражение для дефекта массы, которое переводится в энергию связи  $E_{\text{св}}(Z, A)$  умножением на множитель 931,5 МэВ/а. е. м. Далее находим выражение для удельной энергии связи (в единицах МэВ/нуклон):

$$\frac{E_{\text{св}}(Z, A)}{A} = 15,3 - \frac{16,8}{A^{1/3}} - 87,2 \left( \frac{1}{2} - \frac{Z}{A} \right)^2 - 0,665 \frac{Z^2}{A^{4/3}}. \quad (44.21)$$

Именно эта формула, где вместо  $Z$  использовано выражение (44.17), была графически показана штрихами на рис. 44.1. Модель жидкой капли настолько хорошо передает общий ход кривой, что теоретические и экспериментальные результаты почти неразличимы на рисунке. Разница проявляется в основном в существовании пиков на экспериментальной кривой, наиболее заметным из которых является пик для ядра  ${}^4_2\text{He}$ . Для описания этого явления в формулу для массы и энергии связи вводят еще одно слагаемое, которое мы здесь обсуждать не будем.

## Оболочечная модель ядра

Кривая удельной энергии связи на рис. 44.1 не идет гладко, а имеет пики, соответствующие наиболее устойчивым ядрам. Первый из них соответствует ядру  ${}^4_2\text{He}$ . Опыт показывает, что есть и другие наиболее устойчивые ядра, получившие название *магических*. У магических ядер число протонов  $Z$  или число нейтронов  $N$  равно *магическим числам*: 2, 8, 20, 28, 50, 82, 126. К магическим ядрам по числу протонов относятся, скажем,  ${}^{58}_{28}\text{Ni}$ ,  ${}^{118}_{50}\text{Sn}$ , по числу нейтронов —  ${}^{39}_{19}\text{K}$ ,  ${}^{52}_{24}\text{Cr}$ . Например, энергия связи нейтрона в  ${}^{16}\text{O}$  равна 15,7 МэВ, а в ядре  ${}^{17}\text{O}$  —

всего лишь 4,1 МэВ. У элементов с магическим числом протонов имеется большее число стабильных и долгоживущих изотопов (например, у олова Sn есть десять устойчивых изотопов, а у соседей олова по таблице Менделеева их гораздо меньше: у индия In — один, у сурьмы Sb — два). Ядра, у которых магическими являются  $Z$ , и  $N$ , называются *дважды магическими*. В природе существует всего пять дважды магических ядер: гелий ( ${}^4_2\text{He}$ ), кислород ( ${}^{16}_8\text{O}$ ), два изотопа кальция ( ${}^{40}_{20}\text{Ca}$  и  ${}^{48}_{20}\text{Ca}$ ) и свинец ( ${}^{208}_{82}\text{Pb}$ ).

Появление пиков на кривой удельной энергии связи, существование некоего ряда целых чисел, играющих какую-то роль в свойствах ядер, вызывает в памяти аналогии с атомом. Возникает идея, что наиболее устойчивые ядра, подобно атомам инертных газов, имеют заполненную оболочку из нуклонов. Эту аналогию развивает *оболочечная модель ядра*. Согласно этой модели, каждый нуклон движется в общем ядерном потенциале, создаваемом всеми нуклонами ядра. Этот усредненный потенциал можно аппроксимировать трехмерным осциллятором:  $V(r) = -V_0 + m\omega^2 r^2/2$ . Уровни энергии определяются как сумма энергий трех независимых одномерных осцилляторов:

$$E = -V_0 + \hbar\omega \left( n + \frac{3}{2} \right), \quad n = n_1 + n_2 + n_3, \quad n_i \geq 0. \quad (44.22)$$

Здесь  $n_1, n_2, n_3$  — квантовые числа колебаний нуклона вдоль одной из трех пространственных осей. Эти числа могут быть различными, но энергия зависит от их суммы  $n$ , поэтому возможны разные состояния нуклонов с той же энергией — вырождение уровней, связанное с равноправием координатных осей  $x, y, z$ .

**Задача 44.67.** Определить кратность вырождения уровней энергии трехмерного осциллятора.

**Решение.** При заданном  $n$  возможны различные варианты значений квантовых чисел  $n_1, n_2, n_3$ . Число этих вариантов и есть искомая кратность вырождения  $K_n$ . Первое квантовое число  $n_1$  может принимать любое целое значение от 0 до  $n$ . При каком-то из этих значений сумма двух других квантовых чисел равна  $n_2 + n_3 = n - n_1$ . Ясно, что число возможных вариантов при фиксированных  $n$  и  $n_1$  равно числу значений, которые принимает  $n_2$ . Так как оно может меняться от 0 до  $n - n_1$ , это число равно  $n - n_1 + 1$ . Суммируя по всем возможным значениям  $n_1$ , находим кратность вырождения  $K_n$ :

$$K_n = \sum_{n_1=0}^n (n - n_1 + 1).$$

Вводя новую переменную суммирования  $m = n - n_1 + 1$ , получаем

$$K_n = \sum_{m=1}^{n+1} m = \frac{(n+1)(n+2)}{2}. \quad (44.23)$$

Итак, для основного состояния  $n = 0$  имеем  $K_0 = 1$ , т.е. оно не вырождено. При  $n = 1$  получаем из (44.23)  $K_1 = 3$ . Первое возбужденное состояние трехкратно вырождено, так как данному значению энергии соответствуют следующие тройки квантовых чисел  $(n_1 n_2 n_3)$ : (100), (010) и (001). Второе возбужденное состояние вырождено шестикратно ( $K_2 = 6$ ), так как такую энергию имеют состояния с наборами квантовых чисел  $(n_1 n_2 n_3)$ : (200), (020), (002), (110), (101), (011). ■

Состояния с данным  $n$  называются *оболочками*. Найдем теперь, для каких ядер эти оболочки будут совершенно заполнены. По аналогии с атомом мы можем

**Таблица 44.1.** Номер последней заполненной оболочки  $K$  (первая строка) и общее число  $Z(K)$  протонов (нейтронов) на оболочках с номерами  $1, 2, \dots, K$  в осцилляционном потенциале (вторая строка) в сравнении с магическими числами  $Z_{\text{маг}}$  протонов и нейтронов в ядрах (третья строка).

$K$	1	2	3		4	5	6
$Z(K)$	2	8	20		40	70	112
$Z_{\text{маг}}$	2	8	20	28	50	80	126

ожидать, что таким путем мы получим магические числа протонов и нейтронов. На каждом уровне могут, в соответствии с принципом Паули, помещаться два протона и два нейтрона. Максимальное число протонов (или нейтронов) на данной оболочке равно  $2K_n$ . Например, первая оболочка ( $n = 0$ ) вмещает 2 протона и 2 нейтрона, и она заполняется для элемента  ${}^4_2\text{He}$ . На второй оболочке ( $n = 1$ ) помещается еще 6 протонов и 6 нейтронов; обе оболочки заполняются, таким образом, в ядре  ${}^{16}_8\text{O}$ .

Выведем теперь общую формулу. Число протонов  $Z(K)$ , помещающихся на первых  $K$  оболочках, дается тогда выражением

$$Z(K) = \sum_{n=0}^{K-1} 2K_n = \sum_{n=0}^{K-1} (n^2 + 3n + 2).$$

Используя известные результаты

$$\sum_{n=0}^{K-1} n^0 = K, \quad \sum_{n=0}^{K-1} n = \frac{(K-1)K}{2}, \quad \sum_{n=0}^{K-1} n^2 = \frac{(K-1)K(2K-1)}{6},$$

находим:

$$Z(K) = \frac{K(K+1)(K+2)}{3}. \quad (44.24)$$

Эта формула справедлива также и для числа нейтронов.

В таблице 44.1 приведены результаты расчета  $Z(K)$  по формуле (44.24) в сравнении с магическими значениями, взятыми из эксперимента.

Видно, что первые три магических числа воспроизведены такой простой моделью, но уже четвертое (28) пропущено, а в последующих наблюдаются сильные расхождения с экспериментом. Учет энергии взаимодействия спина нуклона с его орбитальным моментом импульса усложняет модель, но позволяет объяснить весь ряд магических чисел. Модель оболочек объясняет также спины основных и слабо возбужденных состояний ядер. В 1963 г. работавшие независимо американский физик Мария Гёпферт-Майер<sup>1</sup> и немецкий физик Г. Йенсен получили Нобелевскую премию «за открытия, касающиеся оболочечной структуры ядра».

<sup>1</sup>М. Гёпферт-Майер — одна из двух женщин, получивших Нобелевскую премию по физике. Первой была Мария Склодовская-Кюри, ставшая лауреатом этой премии ровно за 60 лет до Гёпферт-Майер.

Мы коснулись лишь двух моделей ядра, исторически бывших самыми первыми созданиями физиков-теоретиков. Дальнейшее развитие науки вызвало к жизни более сложные теории ядер, на которых мы не будем останавливаться.

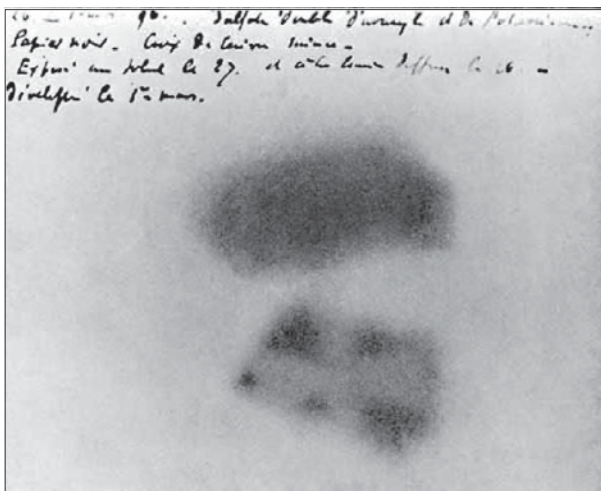
## 44.5 Радиоактивность

♦ *Самопроизвольное превращение одних атомных ядер в другие, которое сопровождается испусканием элементарных частиц, ядер и жесткого электромагнитного излучения, называется **радиоактивностью**.* ♦

Естественная радиоактивность, наблюдаемая у существующих в природе ядер, была открыта в 1896 г. французским ученым А. Беккерелем. Как часто бывает, открытие было сделано случайно, и, как случается не менее часто, в подобных случайностях можно увидеть интуитивную прозорливость. Беккерель происходил из семьи, давшей четыре поколения ученых. Узнав об открытии Рентгеном X-лучей (ноябрь 1895 г.), статья которого сопровождалась впечатляющим снимком руки известного швейцарского анатома и зоолога А. фон Кёлликера (рис. 44.2), и руководствуясь соображениями, что это излучение может быть получено не только в катодной трубке, но и от других источников, Беккерель занялся исследованием излучения образцов, возникающего под действием солнечного света. Он вкладывал фотопластинку между листами плотной черной бумаги, сверху помещал исследуемые образцы (ими были соли урана) и выставлял на солнечный свет. После проявления пластинка оказывалась засвеченной. Если между конвертом с фотопластинкой и солями урана помещался металлический предмет, на проявленной фотопластинке было отчетливо видно его теневое изображение (рис. 44.3). Однажды, в конце февраля 1896 г., опыт пришлось отложить из-за пасмурной погоды,



**Рис. 44.2.** Первый в мире рентгеновский снимок, сделанный самим Рентгеном в январе 1896 г.



**Рис. 44.3.** Открытие Беккерелем радиоактивности солей урана в феврале 1896 г.: отчетливо видно изображение металлического мальтийского креста, помещенного между фотопластинкой и радиоактивным образцом солей урана

и подготовленная фотопластинка лежала в полутемной лаборатории, так что она не подвергалась экспозиции под солнечными лучами. Через несколько дней Беккерель все же проявил пластинку, ожидая увидеть очень слабые следы излучения, но, к своему удивлению, обнаружил на ней интенсивное изображение. Проведя серию опытов с разными веществами, Беккерель уже в мае 1896 г. пришел к правильному выводу, что невидимое излучение испускается самими солями урана и существует вне всякой связи с облучением их солнечным светом. В 1903 г. Беккерель получил Нобелевскую премию «в знак признания исключительных услуг, которые он оказал науке своим открытием самопроизвольной радиоактивности». Вторая половина этой премии была вручена супругам Пьеру и Марии Кюри «в знак признания исключительных услуг, которые они оказали науке совместными исследованиями явлений радиации, открытой профессором Анри Беккерелем»<sup>2</sup>.

Процессы радиоактивных превращений включают:

- $\alpha$ -распад, при котором возникают  $\alpha$ -частицы — ядра гелия;
- $\beta$ -распад (электронный и позитронный, а также и захват ядром электронов); именно в таких процессах были обнаружены очень легкие и не имеющие электрического заряда (нейтральные) частицы нейтрино  $\nu_e$ , появляющиеся вместе с позитронами  $e^+$  (античастицами электрона), и антинейтрино  $\bar{\nu}_e$ , возникающие вместе с электронами;
- $\gamma$ -излучение ядер;
- спонтанное деление тяжелых ядер;
- протонную радиоактивность.

О некоторых из них мы поговорим подробнее чуть позже, но пока отметим, что физические механизмы этих распадов различны, но общим является то, что радиоактивное ядро неустойчиво, т. е. не может жить вечно, и число радиоактивных ядер со временем уменьшается.

## Закон радиоактивного распада

Закон радиоактивного распада определяет количество нераспавшихся атомов в любой момент времени. Пусть в момент  $t$  имеется  $N$  радиоактивных ядер. Учитывая, что ядра претерпевают превращения независимо друг от друга, можно считать, что количество  $dN$  распадающихся за интервал  $dt$  ядер пропорционально их общему числу и интервалу времени, т. е.

$$dN = -\lambda N dt. \quad (44.25)$$

Знак «минус» свидетельствует об уменьшении числа ядер вследствие происходящего их распада. Коэффициент пропорциональности  $\lambda$  называется *постоянной*

<sup>2</sup>Заинтересовавшись открытием Беккереля, Мария Кюри, выпускница знаменитой Сорбонны — Парижского университета, выбрала исследование радиоактивности своим научным направлением. Вместе с Пьером Кюри, известным своими исследованиями зависимости магнитных свойств веществ от температуры (о *точке Кюри* см. т. 1, разд. 21.9) и пьезоэлектрического эффекта, они в 1898 г. открыли два радиоактивных элемента — полоний (названный так в честь Польши, родины Марии Склодовской-Кюри) и радий. В 1911 г., уже после нелепой гибели мужа под колесами конной повозки, Мария Кюри была удостоена второй Нобелевской премии — на этот раз по химии — «за выдающиеся заслуги в развитии химии: открытие элементов радия и полония, выделение радия и изучение природы и соединений этого замечательного элемента».

**радиоактивного распада.** Интегрирование дифференциального уравнения (44.25) дает **закон радиоактивного распада**

$$\boxed{N(t) = N_0 e^{-\lambda t}}, \quad (44.26)$$

где  $N_0$  — количество ядер в начальный момент времени, т. е.  $N_0 = N(0)$ , а  $N(t)$  — число нераспавшихся ядер в момент времени  $t$ .

Физический смысл постоянной распада  $\lambda$  становится ясным из следующих рассуждений. В начальный момент времени имеется  $N_0$  ядер, к моменту времени  $t$  их число равно  $N(t)$ . В интервале  $(t, t + dt)$  распадается  $-dN$  ядер. Поэтому отношение  $-dN/N_0$  есть вероятность того, что ядро «доживет» до момента  $t$ , но не «переживет» момента  $t + dt$ . Иными словами, это вероятность того, что ядро «умрет» в возрасте  $t$ . Умножая вероятность на возраст  $t$  ядра и интегрируя по всем возможным временам, получаем отсюда среднее время жизни ядра:

$$\tau = -\frac{1}{N_0} \int_0^{\infty} t dN = \frac{\lambda}{N_0} \int_0^{\infty} t N(t) dt = \lambda \int_0^{\infty} t e^{-\lambda t} dt = \frac{1}{\lambda}. \quad (44.27)$$

Таким образом, постоянная распада обратна среднему времени жизни ядра.

◆ **Периодом полураспада**  $T_{1/2}$  называется время, в течение которого распадается половина первоначального количества ядер. ◆

Для нахождения  $T_{1/2}$  получаем из (44.26) уравнение

$$\frac{N_0}{2} = N_0 e^{-\lambda T_{1/2}}, \quad (44.28)$$

откуда

$$\boxed{T_{1/2} = \frac{\ln 2}{\lambda} = \frac{0,693}{\lambda} = 0,693\tau.} \quad (44.29)$$

Период полураспада лишь численным множителем отличается от среднего времени жизни ядра.

Нередки случаи, когда радиоактивный элемент имеет два или более каналов распада с разными постоянными распада  $\lambda_i$ . Тогда  $\lambda_i N dt$  — число ядер, распавшихся за время  $dt$  по каналу с номером  $i$ , и изменение числа изначальных «материнских» ядер равно

$$dN = -N \sum_i \lambda_i dt = -\lambda N dt, \quad \lambda = \sum_i \lambda_i, \quad (44.30)$$

где введена эффективная постоянная  $\lambda$  распада ядра по всем каналам. Вероятность  $p_i$  распада ядра по каналу  $i$  равна отношению распавшихся по этому каналу ядер к общему числу распадов за это же время:

$$p_i = \frac{\lambda_i N dt}{\lambda N dt} = \frac{\lambda_i}{\lambda}, \quad (44.31)$$

откуда можно найти постоянные распада  $\lambda_i = p_i \lambda$ . Если ввести *парциальный период полураспада*  $T_{1/2}^{(i)}$  по каналу  $i$ , т. е. период полураспада, который был бы у данной системы, если «выключить» все другие каналы распада, кроме  $i$ -го, то получим соотношения

$$T_{1/2}^{(i)} = \frac{\ln 2}{\lambda_i} = \frac{\ln 2}{p_i \lambda} = \frac{T_{1/2}}{p_i}. \quad (44.32)$$

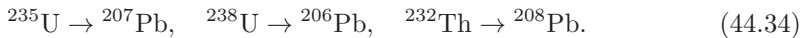
**Задача 44.68.** Радиоактивный изотоп калий-40 в 89,28% случаев претерпевает  $\beta$ -распад  ${}^{40}_{19}\text{K} \rightarrow {}^{40}_{20}\text{Ca} + e^- + \bar{\nu}_e$ , а в 10,72% случаев захватывает орбитальный электрон, превращаясь в ядро аргона-40 в результате ядерной реакции (так называемого К-захвата)  ${}^{40}_{19}\text{K} + e^- \rightarrow {}^{40}_{18}\text{Ar} + \nu_e + \gamma$ . Период полураспада калия-40 равен  $T_{1/2} = 1,251$  млрд лет. Найти парциальные периоды полураспада в кальций и аргон.

**Решение.** Вероятности распада по этим каналам равны  $p_{\text{Ca}} = 0,8928$  и  $p_{\text{Ar}} = 0,1072$ . Из формулы (44.32) тогда следуют результаты

$$T_{1/2}^{(\text{Ca})} = \frac{1,251}{0,8928} = 1,40 \text{ млрд лет}, \quad (44.33)$$

$$T_{1/2}^{(\text{Ar})} = \frac{1,251}{0,1072} = 11,7 \text{ млрд лет. } \blacksquare$$

В природе также часто случается, что возникающие в результате радиоактивного распада ядра в свою очередь оказываются радиоактивными. Следовательно, возникает целый ряд радиоактивных превращений (цепочка распадов). Существуют три **радиоактивных семейства**, родоначальниками которых являются изотопы урана или тория, а конечными продуктами — изотопы свинца:



Четвертое радиоактивное семейство — нептуниевое — имеет своим родоначальником искусственно полученный в ядерных реакциях изотоп  ${}^{237}_{93}\text{Np}$ .

## Распад «дочерних ядер»

Рассмотрим модельный пример — систему, где продукт распада ядра также оказывается нестабильным. Пусть в замкнутый объем помещено  $N_0$  атомов радиоактивного элемента, постоянная распада для которого равна  $\lambda_1$ . Ядра элемента, образующегося в результате такого распада, также радиоактивны, их постоянная распада равна  $\lambda_2$ .

За время  $dt$  число  $N_2$  ядер нового элемента («дочерних» ядер) изменяется, во-первых, за счет появления новых ядер в результате распада исходных («материнских») и, во-вторых, за счет убыли собственных в результате их распада:

$$dN_2 = \lambda_1 N_1 dt - \lambda_2 N_2 dt. \quad (44.35)$$

Здесь  $N_1$  — число «материнских» ядер, а  $N_2$  — число «дочерних» в тот же момент времени  $t$ . Согласно закону радиоактивного распада  $N_1 = N_0 \exp(-\lambda_1 t)$ , так что

$$\frac{dN_2}{dt} + \lambda_2 N_2 = \lambda_1 N_0 e^{-\lambda_1 t}. \quad (44.36)$$

Решение неоднородного дифференциального уравнения можно представить в виде суммы общего решения однородного уравнения и частного решения неоднородного. Последнее ищем в виде  $Ae^{-\lambda_1 t}$ . Подставляя это выражение в (44.36), получаем соотношение

$$-\lambda_1 A + \lambda_2 A = \lambda_1 N_0, \quad (44.37)$$

откуда

$$A = N_0 \frac{\lambda_1}{\lambda_2 - \lambda_1}. \quad (44.38)$$

Общее решение однородного уравнения  $\dot{N}_2 + \lambda_2 N_2 = 0$  уже хорошо нам известно и может быть представлено в виде  $Be^{-\lambda_2 t}$ . Таким образом,

$$N_2(t) = Be^{-\lambda_2 t} + N_0 \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_1 t}. \tag{44.39}$$

Постоянная  $B$  в этом соотношении определяется из начального условия отсутствия «дочерних» ядер в момент времени  $t = 0$ : соотношение  $N_2(0) = 0$  дает  $B = -N_0 \frac{\lambda_1}{\lambda_2 - \lambda_1}$ . Окончательно получаем

$$N_2(t) = N_0 \frac{\lambda_1}{\lambda_2 - \lambda_1} (e^{-\lambda_1 t} - e^{-\lambda_2 t}). \tag{44.40}$$

В начальный момент и по прошествии достаточно большого времени число «дочерних» ядер равно нулю. В момент  $t_{\max}$  выражение (44.40) достигает своего максимального значения. Величину  $t_{\max}$  можно определить, приравняв нулю производную  $dN_2/dt$ :

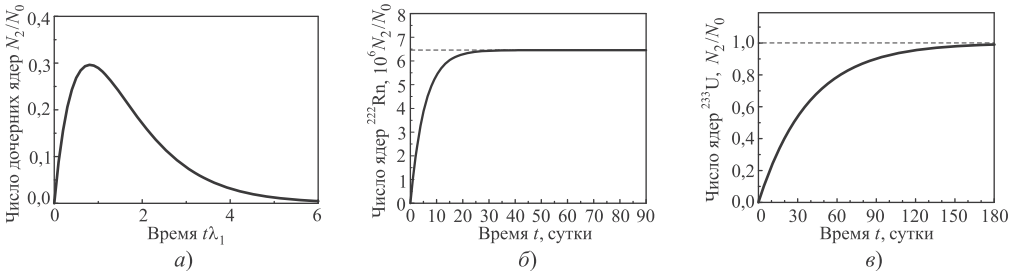
$$t_{\max} = \frac{\ln \lambda_2 - \ln \lambda_1}{\lambda_2 - \lambda_1} = \frac{\ln(\lambda_2/\lambda_1)}{\lambda_2 - \lambda_1}. \tag{44.41}$$

Подставляя (44.41) в (44.40), находим максимальное значение  $N_2$ :

$$N_{2,\max} = N_0 \left( \frac{\lambda_1}{\lambda_2} \right)^{\frac{\lambda_2}{\lambda_2 - \lambda_1}}. \tag{44.42}$$

Общий характер временной зависимости  $N_2(t)$  показан на рис. 44.4, *a* для случая, когда постоянные распада «материнских» и «дочерних» ядер близки. На практике при разумных временах наблюдения показанный пик может быть очень размыт и в сущности ненаблюдаем.

Рассмотрим предельные случаи  $\lambda_1 \ll \lambda_2$  и  $\lambda_1 \gg \lambda_2$ .



**Рис. 44.4.** Временная зависимость относительного числа  $N_2/N_0$  «дочерних» ядер при различных соотношениях между постоянными распада  $\lambda_1$  и  $\lambda_2$  «материнских» и «дочерних» ядер: *a*) при близких значениях постоянных распада  $\lambda_2/\lambda_1 = 1, 5$  отчетливо виден пик в зависимости  $N_2(t)$  (время показано в единицах среднего времени жизни «материнских» ядер  $t\lambda_1$ ); *б*) относительное число ядер изотопа радона  $^{222}_{86}\text{Rn}$ , возникающих при радиоактивном распаде радия  $^{226}_{88}\text{Ra}$  (в этом случае  $\lambda_1 \ll \lambda_2$ ); *в*) относительное число ядер изотопа урана  $^{233}_{92}\text{U}$ , возникающих при радиоактивном распаде протактиния  $^{233}_{91}\text{Pa}$  (здесь, наоборот,  $\lambda_1 \gg \lambda_2$ )



Если  $\lambda_1 \ll \lambda_2$ , то «материнские» ядра распадаются гораздо медленнее «дочерних», и число последних быстро достигает своего максимального значения. Из (44.42) при  $\lambda_2 \rightarrow \infty$  следует связь этого максимального значения с начальным:  $N_{2,\max} \approx N_0 \lambda_1 / \lambda_2$ . Но такое быстрое «подстраивание» числа «дочерних» ядер под число «материнских» происходит в каждый момент времени, так что закон изменения их числа должен иметь вид

$$N_2(t) \approx N_1(t) \frac{\lambda_1}{\lambda_2} = N_0 \frac{\lambda_1}{\lambda_2} e^{-\lambda_1 t}. \quad (44.43)$$

Этот же результат прямо следует из (44.41) при  $\lambda_2 \rightarrow \infty$ . Обращает внимание, что число «дочерних» ядер уменьшается со временем с постоянной распада «материнских» ядер.

Примером может служить радиоактивный распад радия  ${}^{226}_{88}\text{Ra}$  с постоянной распада  $\lambda_1 = 1,354 \cdot 10^{-11} \text{ с}^{-1}$  (период полураспада «материнского» ядра  $T_{1/2}^{(1)} = \ln 2 / \lambda_1 = 1622$  года), продуктом которого является «дочернее» ядро радона  ${}^{222}_{86}\text{Rn}$  с постоянной распада  $\lambda_2 = 2,097 \cdot 10^{-6} \text{ с}^{-1}$  (период полураспада  $T_{1/2}^{(2)} = 3,83$  суток). Если радий поместить в закрытый сосуд, то уже через месяц содержание радона в нем будет лишь на 0,4% меньше равновесного. Само же равновесное значение, как видно из соотношения  $N_2/N_0 = \lambda_1/\lambda_2$ , составляет лишь около шести миллионных исходного числа атомов радия. Достигнутое значение концентрации радона будет сохраняться неизменным столетиями в соответствии с периодом полураспада «материнских» ядер (около 1600 лет). Этот случай показан на рис. 44.4, б. Только на интервалах порядка нескольких тысячелетий станет заметным уменьшение числа атомов радона, хотя период полураспада ядер радона составляет менее четырех суток.

Достаточно часто распад «материнских» ядер идет по нескольким каналам, так что возникают разные «дочерние» радиоактивные продукты, а не лишь один тип ядер. Обычно они все распадаются гораздо быстрее «материнских» ядер. В этом случае полученное нами соотношение  $\lambda_1 N_1 = \lambda_2 N_2$  (см. (44.43)) для числа «материнских» и равновесного числа «дочерних» ядер обобщается очевидным образом:

$$\begin{aligned} \lambda_1 N_1 &= \lambda_2 N_2 = \lambda_3 N_3 = \dots = \lambda_k N_k, \\ \lambda_1 &\ll \lambda_2, \lambda_3, \dots, \lambda_k. \end{aligned} \quad (44.44)$$

В обратном случае  $\lambda_1 \gg \lambda_2$  «материнские» ядра распадаются гораздо быстрее «дочерних», так что уже через несколько средних времен жизни «материнских» ядер они почти исчезают, а число «дочерних» ядер становится почти равным начальному числу «материнских»:  $N_2 \approx N_0$ . Далее «дочерние» ядра распадаются по обычному закону со своей собственной постоянной распада  $\lambda_2$ :  $N_2 = N_0 \exp(-\lambda_2 t)$ .

Примером может служить распад протактиния  ${}^{233}_{91}\text{Pa}$  ( $T_{1/2}^{(1)} = 27$  сут.,  $\lambda_1 = 2,97 \cdot 10^{-7} \text{ с}^{-1}$ ), в результате которого образуется изотоп урана  ${}^{233}_{92}\text{U}$  ( $T_{1/2}^{(2)} = 1,592 \cdot 10^5$  лет,  $\lambda_2 = 1,38 \cdot 10^{-13} \text{ с}^{-1}$ ). Этому случаю соответствует кривая на рис. 44.4, в. В отличие от предыдущего случая равновесная относительная концентрация «дочерних» ядер близка к единице, а ее уменьшение происходит в соответствии с периодом полураспада «дочерних» ядер (урана в нашем примере).

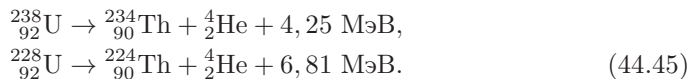
## Альфа-распад

**Альфа-распад** протекает по следующей схеме:  ${}^A_ZX \rightarrow {}^{A-4}_{Z-2}Y + {}^4_2\text{He}$ , где альфа-излучение представляет поток ядер гелия  ${}^4_2\text{He}$ . «Дочернее» ядро  ${}^{A-4}_{Z-2}Y$  обычно оказывается возбужденным и теряет энергию возбуждения, испуская  $\gamma$ -кванты (жесткое электромагнитное излучение). Ядра гелия  ${}^4_2\text{He}$  ( $\alpha$ -частицы в исторически сложившейся терминологии, возникшей еще до выяснения природы различных типов излучения) вылетают из «материнского» ядра  ${}^A_ZX$  с большими скоростями порядка  $10^7$  м/с (т. е.  $0,1c$ ), а их кинетическая энергия составляет от 4 до 10 МэВ. Кинетическая энергия  $\alpha$ -частиц возникает за счет дефекта массы «материнского» ядра по отношению к «дочернему» ядру и  $\alpha$ -частице. Эта избыточная энергия распределяется между  $\alpha$ -частицей и «дочерним» ядром  $Y$  в отношении, обратном пропорциональному их массам, в соответствии с классическим законом сохранения импульса (т. е.  $E_\alpha : E_Y = m_Y : m_\alpha$ ). Поэтому энергия альфа-частиц при распаде любых ядер данного изотопа всегда одна и та же.

Если, как это обычно бывает, масса «дочернего» ядра намного превышает массу  $\alpha$ -частицы, то почти вся выделяемая энергия будет унесена в виде кинетической энергии  $\alpha$ -частицы.

Возможно испускание не одной, а нескольких групп моноэнергетических альфа-частиц, соответствующих различным значениям энергии «дочерних» ядер в разных возбужденных состояниях.

Двигаясь в веществе,  $\alpha$ -частицы испытывают неупругие соударения, постепенно теряют свою энергию (затрачивая ее на ионизацию молекул вещества) и останавливаются. Пробег  $\alpha$ -частиц до остановки существенно зависит от плотности вещества. Например, в воздухе при нормальных условиях их пробег имеет величину порядка десятка микрометров (другими словами, обычный лист бумаги полностью экранирует  $\alpha$ -излучение). Примером альфа-распада могут служить распад природного урана с образованием тория и аналогичный распад радиоактивного изотопа урана:



Мы выбрали эти распады в качестве примера для дальнейшего обсуждения, поскольку, несмотря на похожесть, они резко отличаются скоростями протекания реакции.

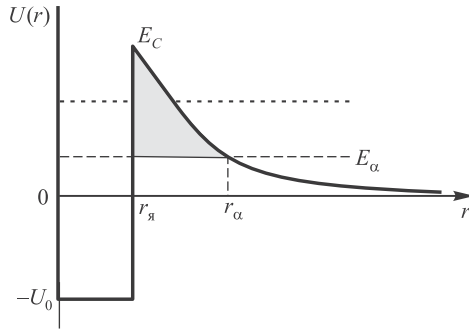
**Задача 44.69.** На примере распада природного урана оценить энергию кулоновского взаимодействия  $\alpha$ -частицы и «дочернего» ядра в момент ее вылета из «материнского» ядра.

**Решение.** Радиус «материнского» ядра  ${}^A_ZX$  равен  $r_X = r_p A^{1/3}$ . Именно на таком расстоянии находится  $\alpha$ -частица в момент вылета. Заряд «дочернего» ядра равен  $q_d = (Z - 2)e$ , заряд  $\alpha$ -частицы равен  $q_\alpha = 2e$ , т. е. энергия кулоновского отталкивания равна

$$E_C = \frac{e^2 2(Z - 2)}{4\pi\epsilon_0 r_p A^{1/3}} = E_p \frac{2(Z - 2)}{A^{1/3}}.$$

Входящий сюда множитель  $E_p$  одинаков для всех ядер:

$$E_p = \frac{e^2}{4\pi\epsilon_0 r_p} = 1,108 \text{ МэВ}. \quad (44.46)$$



**Рис. 44.5.** Туннелирование  $\alpha$ -частицы при распаде: штриховой прямой показана энергия  $E_\alpha$  альфа-частицы; область под барьером, недоступная частице с точки зрения классической механики, показана темным фоном; при том же кулоновском потенциале  $U$ , но большей энергии  $E_\alpha$  (пунктир) высота и ширина барьера уменьшаются, что делает туннелирование гораздо более вероятным, как в случае распада  $^{228}\text{U}$  по сравнению с распадом  $^{238}\text{U}$

Теперь наша формула для энергии кулоновского взаимодействия может быть записана в виде

$$E_C = 1,108 \text{ МэВ} \frac{2(Z-2)}{A^{1/3}}. \quad (44.47)$$

Подставляя в (44.47) значения  $A = 238$ ,  $Z = 92$  для урана, получаем

$$E_C \approx 1,108 \frac{2 \times 90}{238^{1/3}} \approx 32,2 \text{ МэВ}. \quad (44.48)$$

Близкий результат  $E_C = 32,6 \text{ МэВ}$  получается и для ядра  $^{228}\text{U}$ . ■

## Элементы теории альфа-распада

Исходя из полученной оценки, мы можем представить себе потенциал взаимодействия  $\alpha$  частицы с ядром в следующем виде. На расстояниях  $r \leq r_\text{я}$   $\alpha$ -частица притягивается к ядру огромными, но короткодействующими ядерными силами. Мы описываем это взаимодействие простейшим потенциалом прямоугольной формы, учитывающим общее влияние всех нуклонов ядра. Внутри этой потенциальной ямы  $\alpha$ -частица движется свободно. Со стороны протонов на заряженную альфа-частицу действуют также силы кулоновского отталкивания. Эти силы значительно слабее ядерных, но убывают с расстоянием очень медленно (дальнодействующие силы). На расстояниях же  $r \geq r_\text{я}$  ядерные силы уже перестают действовать, и  $\alpha$ -частица отталкивается от ядра кулоновскими силами. Таким образом, потенциал имеет вид (рис. 44.5)

$$V(r) = \begin{cases} -V_0, & r < r_\text{я}, \\ E_C \frac{r_\text{я}}{r}, & r > r_\text{я}. \end{cases} \quad (44.49)$$

Если  $\alpha$ -частица вылетела из ядра с энергией  $E_\alpha = 4,2 \text{ МэВ}$ , то такова была ее энергия на основном уровне в ядерной части потенциала. Как же она смогла преодолеть высокий кулоновский барьер и вылететь наружу? Здесь проявился изученный нами ранее квантовомеханический туннельный эффект.

Туннельный механизм альфа-распада позволяет оценить время жизни «материнского» ядра. Пусть альфа-частица совершает внутри ядра колебательное движение со

скоростью  $v$ . Время между двумя последовательными столкновениями альфа-частицы с барьером равно  $\Delta t = 2r_\alpha/v$ . Пусть коэффициент прозрачности барьера равен  $D$ . Эта величина дает вероятность прохождения через барьер. За время  $\tau$  альфа-частица ударится о барьер  $\tau/\Delta t$  раз, так что вероятность вылететь наружу за время  $\tau$  составит  $D\tau/\Delta t$ . Приравнявая эту вероятность единице, получаем время жизни радиоактивного ядра:

$$\tau = \frac{\Delta t}{D} = \frac{2r_\alpha}{vD}. \quad (44.50)$$

В соответствии с формулой (36.41) можно вычислить коэффициент прозрачности  $D$  кулоновского барьера. Подставляя в это выражение кулоновскую часть потенциала (44.49), получаем:

$$D = \exp \left[ -\frac{2}{\hbar c} \sqrt{2E_0} \int_{r_\alpha}^{r_\alpha} dr \sqrt{E_C \frac{r_\alpha}{r} - E_\alpha} \right], \quad (44.51)$$

где  $E_0 = m_\alpha c^2 = 3728$  МэВ — энергия покоя  $\alpha$ -частицы, а  $r_\alpha = r_\alpha(E_C/E_\alpha)$  — расстояние, при котором энергия  $\alpha$ -частицы равна энергии кулоновского отталкивания (см. рис. 44.5). Выполнив в интеграле замену переменных  $r = x^2 r_\alpha E_C/E_\alpha$ , переписываем (44.51) в виде:

$$D = \exp \left\{ -\frac{4}{\hbar c} r_\alpha E_C \sqrt{\frac{2E_0}{E_\alpha}} \int_{\sqrt{E_\alpha/E_C}}^1 dx \sqrt{1-x^2} \right\} = \\ = \exp \left[ -\frac{2}{\hbar c} r_\alpha E_C \sqrt{\frac{2E_0}{E_\alpha}} \left( \arccos \sqrt{\frac{E_\alpha}{E_C}} - \sqrt{\frac{E_\alpha}{E_C}} \sqrt{1 - \frac{E_\alpha}{E_C}} \right) \right]. \quad (44.52)$$

Применим это выражение сначала для  $\alpha$ -распада  $^{238}\text{U}$ . Подставим в (44.52) численные значения:  $r_\alpha E_C = 2(Z-2)r_p E_p = 180 \times 1,3 \times 1,108 = 259,27$  МэВ · фм,  $\hbar c = 197,327$  МэВ · фм,  $E_0 = 3728$  МэВ,  $E_\alpha = 4,25$  МэВ и вычисленное выше значение  $E_C = 32,2$  МэВ. В результате получим  $D_1 = \exp(-94,71) = 7,4 \cdot 10^{-41}$ . Для распада  $^{228}\text{U}$  меняются только значения  $E_\alpha = 6,81$  МэВ и  $E_C = 32,6$  МэВ, при этом прозрачность барьера резко увеличивается:  $D_2 = \exp(-59,96) = 9,1 \cdot 10^{-27}$ .

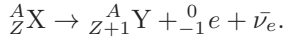
Таким образом, для двух рассмотренных распадов (44.45) отношение соответствующих коэффициентов прозрачности  $D_1$  для  $^{238}\text{U}$  и  $D_2$  для  $^{228}\text{U}$  огромно:  $D_2/D_1 = 1,2 \cdot 10^{14}$ . Обращает внимание чрезвычайная чувствительность коэффициента прозрачности к высоте и ширине барьера: повышение энергии альфа-частицы (при почти одинаковых потенциалах) всего в полтора раза, от 4,25 МэВ до 6,81 МэВ, и вызванное этим уменьшение барьера приводит к увеличению коэффициента прозрачности на 14 порядков! Соответственно второй распад будет происходить гораздо быстрее первого. Поскольку расхождение в остальных параметрах (скорость альфа-частиц, радиусы ядер) для этих ядер не столь велико, отношение времен жизни данных изотопов  $\tau_1/\tau_2 \approx D_2/D_1 = 1,2 \cdot 10^{14}$ . Эксперимент дает следующие значения периодов полураспада:  $T_{1/2}^{(1)} = 4,9 \cdot 10^9$  лет =  $1,55 \cdot 10^{17}$  с,  $T_{1/2}^{(2)} = 9,1$  мин = 546 с, откуда следует отношение  $T_{1/2}^{(1)}/T_{1/2}^{(2)} = 3 \cdot 10^{14}$ . Это совпадает по порядку величины с данной нами оценкой, что указывает на правильность объяснения альфа-распада в рамках механизма квантового туннелирования. Становятся понятными главные особенности альфа-распада:

- постоянство энергии  $\alpha$ -частиц при распаде ядер данного типа;
- малый разброс энергий при сравнении распадов разных ядер;
- широкий разброс периодов полураспада для различных ядер.

## Бета-распад

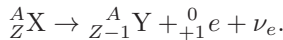
**Бета-распад** ядер связан с испусканием ядром *электрона* (или его античастицы с положительным зарядом — *позитрона*). К этому же типу процессов относится *электронный K-захват*, когда ядро поглощает один из электронов атома (в основном с K-оболочки).

**Электронный распад ( $\beta^-$ -распад)** происходит по схеме



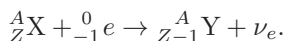
Здесь для электрона (массовое число 0, зарядовое число  $-1$ ) использованы «ядерные» символы. Вместе с электроном появляется еще одна частица — *антинейтрино*, для которой массовое и зарядовое числа равны нулю. Процесс протекает так, как если бы один из нейтронов «материнского» ядра  ${}^A_Z X$  превратился бы в протон с испусканием электрона и антинейтрино ( $n \rightarrow p + e^- + \bar{\nu}_e$ ). При этом «дочернее» ядро может возникнуть как в нормальном (невозбужденном) состоянии, так и (чаще) в возбужденном. В таком случае образовавшееся ядро переходит в состояние с меньшей энергией, при этом «высвечивается» фотон, т. е. бета-распад обычно сопровождается  $\gamma$ -излучением. Примером  $\beta^-$ -распада может служить превращение изотопа углерода (*радиоуглерода*) в азот:  ${}^{14}_6 C \rightarrow {}^{14}_7 N + {}^0_{-1} e + \bar{\nu}_e$ .

**Позитронный распад ( $\beta^+$ -распад)** происходит по схеме

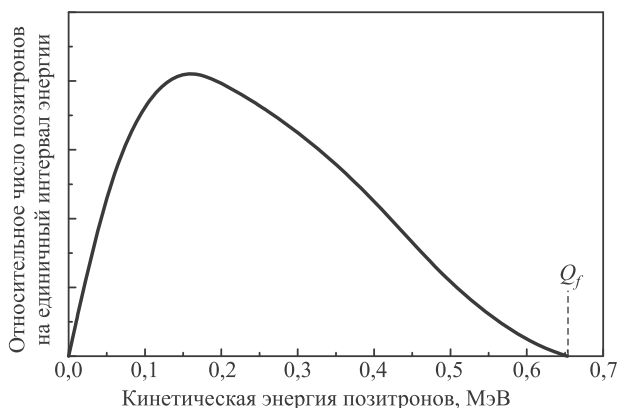


Процесс выглядит как распад протона ( $p \rightarrow n + e^+ + \nu_e$ ). Эта реакция получается из распада нейтрона применением двух правил: 1) если возможен некий процесс, то возможен и обратный, т. е. можно обратить стрелку в реакции, и 2) при переносе частицы из одной части уравнения в другую надо сменить все ее «заряды», т. е. заменить частицу на античастицу. При таком переносе электрон у нас заменился на позитрон, а антинейтрино — на нейтрино. Заметим, что в свободном состоянии протон распадаться не может, так как его масса меньше массы нейтрона. Однако в ядре важен баланс масс «материнского» и «дочернего» ядер, и подобные процессы возможны за счет изменения энергии связи нуклонов. В качестве примера укажем превращение меди в никель:  ${}^{64}_{29} Cu \rightarrow {}^{64}_{28} Ni + {}^0_{+1} e + \nu_e$ .

**Электронный захват**, или **K-захват** — третий тип  $\beta$ -распада — происходит при поглощении ядром одного из электронов K-оболочки, в результате чего один из протонов превращается в нейтрон с испусканием нейтрино:  $p + e^- \rightarrow n + \nu_e$ . Эта реакция получается из распада протона (связанного в ядре) перенесением в левую часть позитрона с заменой его на электрон. На уровне ядра схема электронного захвата записывается в виде



В электронной K-оболочке «материнского» атома образуется вакантное место, которое заполняется электроном из вышележащих энергетических состояний, в результате чего возникает каскад рентгеновского излучения. Электронный захват происходит при превращении калия в аргон:  ${}^{40}_{19} K + {}^0_{-1} e \rightarrow {}^{40}_{18} Ar + \nu_e$ .



**Рис. 44.6.** Распределение энергий позитронов, испускаемых при распаде  ${}^{64}_{29}\text{Cu} \rightarrow {}^{64}_{28}\text{Ni} + {}^0_{+1}e + \nu_e$ . Предельное значение энергии позитрона  $Q_f = 0,653$  МэВ. Видно, что чаще всего будут наблюдаться позитроны с энергией  $E_e \sim 0,15$  МэВ, и в этих случаях на долю нейтрино приходится энергия  $E_{\nu_e} \sim 0,5$  МэВ.

Важно отметить, что при всех видах радиоактивности выполняются законы сохранения массового числа, зарядового числа, а также *лептонного числа*<sup>3</sup>.

Может показаться удивительным, что ядра испускают электроны и нейтрино, а мы все время утверждали, что они состоят из нуклонов. Однако вспомним, что атомы испускают фотоны, а фотоны не содержатся в атомах, но порождаются при переходе электрона с вышележащего уровня на нижний. Электроны и нейтрино не содержатся в ядрах, но рождаются при превращениях протонов в нейтроны и обратно.

Как и при альфа-распаде, в рассмотренных процессах выделяется строго фиксированное количество энергии  $Q_f$ , определяемое дефектом массы «материнского» и «дочернего» ядер. Однако при бета-распаде эта энергия делится между позитроном (электроном) и нейтрино (антинейтрино). Иногда почти всю энергию уносит электрон, иногда — нейтрино. Вероятность того или иного распределения энергии определяется законами данного типа взаимодействий (их называют *слабыми взаимодействиями*). Поэтому при измерении энергетического спектра электронов (позитронов) получается не строго фиксированная энергия, а некоторое распределение в диапазоне от нуля (когда вся энергия уносится нейтрино) до порогового значения  $Q_f$  (если вся выделенная энергия приходится на электрон). Примерный вид энергетического спектра продуктов бета-распада показан на рис. 44.6. В свое время подобные экспериментальные результаты вызвали замешательство среди физиков, поскольку, как им тогда

<sup>3</sup>Физический смысл лептонного числа, приписываемого некоторым частицам (*лептонам*, о которых см. подробнее в следующей главе), остается неясным, но его формальное введение в теорию «объясняет», почему электрон рождается в паре с антинейтрино, а позитрон — в паре с нейтрино. Здесь, как и во многих других случаях, физики следуют совету Мефистофеля из трагедии И. В. Гёте «Фауст»:

*Коль скоро недочет в понятиях случится,  
Их можно словом заменить.*

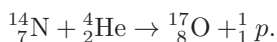
казалось, эти результаты указывали на нарушение закона сохранения энергии. В 1930 г. швейцарский физик В. Паули предположил, что недостающая часть энергии уносится невидимой нейтральной частицей — нейтрино (название придумано Э. Ферми — уменьшительное от «нейтрон»). Прямое наблюдение потока нейтрино от атомного реактора было осуществлено в 1953 г. Ф. Райнесом и К. Коуэном в США. Это открытие принесло Райнесу в 1995 г. Нобелевскую премию «за экспериментальное обнаружение нейтрино» (Коуэн к этому времени уже скончался, а эти премии посмертно не присуждаются). Ну а Паули получил свою Нобелевскую премию еще в 1945 г. «за открытие принципа запрета, названного его именем».

## Ядерные реакции и определение возраста археологических образцов радиоуглеродным методом

♦ *Различные процессы сильного взаимодействия атомных ядер с элементарными частицами или с другими ядрами носят название ядерных реакций.* ♦

Ядерные реакции возникают при сближении сильно взаимодействующих частиц до малых расстояний порядка  $10^{-15}$  м. Неудачи древних алхимиков по превращению элементов связаны с тем, что они пытались использовать химические реакции, для которых характерны энергии порядка нескольких электронвольт, в то время как энергии в ядерных процессах в миллион раз больше. В настоящее время с помощью современных ускорителей удается даже получать новые трансурановые элементы, отсутствующие в естественном состоянии. Подобными работами прославлен Объединенный институт ядерных исследований (ОИЯИ) в Дубне, и признание международного научного сообщества отразилось в официальных именах, присвоенных открытым там трансурановым элементам<sup>4</sup>.

Наиболее распространены ядерные реакции, при которых легкая частица  $a$  сталкивается с ядром  $X$ , в результате чего образуется новое ядро  $Y$  и легкая частица  $b$ :  $X + a \rightarrow Y + b$ . Уравнения таких реакций принято также записывать в сокращенном виде  $X(a, b)Y$ , где в скобках указываются сначала исходная легкая частица, а потом через запятую — конечная. Впервые ядерная реакция была осуществлена в 1919 г. Э. Резерфордом. При облучении азота  $\alpha$ -частицами некоторые ядра азота превращались в ядра кислорода, испуская при этом протон. Уравнение этой реакции имеет вид  ${}^{14}_7\text{N}(\alpha, p){}^{17}_8\text{O}$ , или, в более привычной записи:



На самом деле мы уже неоднократно выписывали уравнения различных ядерных превращений, и настало время подчеркнуть то, что читатель, быть может, уже и сам заметил: суммы верхних (нижних) индексов справа и слева от стрелки совпадают. Суммы верхних индексов дают *число нуклонов, которое сохраняется*

<sup>4</sup> Сто пятый элемент *дубний*  ${}^{262}_{105}\text{Db}$  увековечил имя маленькой подмосковной Дубны. *Флеровий*  ${}^{289}_{114}\text{Fl}$  назван в честь Лаборатории ядерных реакций (ЛЯР) ОИЯИ и ее основателя и директора Г. Н. Флёрва (1913–1990). *Московский*  ${}^{288}_{115}\text{Mc}$  назван в честь Московской области, где расположена Дубна. *Оганесон*  ${}^{294}_{118}\text{Og}$  получил имя в честь Ю. Ц. Оганесяна, директора ЛЯР им. Г. Н. Флёрва в 1989–1996 гг., а ныне научного руководителя этой Лаборатории.

в реакции. Равенство сумм нижних индексов означает закон сохранения электрического заряда.

Особый интерес представляет ядерная реакция  ${}^{14}_7\text{N}(n, p){}^{14}_6\text{C}$ , которая происходит в стратосфере на высотах 12–15 км под действием нейтронов, образуемых космическим излучением. При этом возникает так называемый радиоуглерод  ${}^{14}_6\text{C}$  — бета-радиоактивный изотоп с периодом полураспада  $T_{1/2} = 5730 \pm 40$  лет (распад идет по схеме  ${}^{14}_6\text{C} \rightarrow {}^{14}_7\text{N} + {}^0_{-1}e + \bar{\nu}_e$ , максимальная энергия электронов  $Q_f = 0,156$  МэВ). По химическим свойствам радиоуглерод ничем не отличается от обычного углерода, который представляет собой смесь стабильных изотопов  ${}^{12}\text{C}$  (98,89%) и  ${}^{13}\text{C}$  (1,11%). Убытие радиоуглерода вследствие распада компенсируется его воспроизводством под действием космических лучей, так что в целом его содержание поддерживается на некотором равновесном уровне. Исследования показали, что в каждом грамме природного углерода в минуту происходит примерно 14 распадов радиоуглерода. Из атмосферы посредством обмена радиоуглерод переходит в гидросферу и биосферу. Соединения углерода являются основой жизни на Земле. В процессе обмена веществ (фотосинтеза) живые организмы усваивают из воздуха углекислоту  $\text{CO}_2$ , часть молекул которой оказывается, таким образом, радиоактивной. Концентрация радиоуглерода в живом организме (растении) равна его концентрации в атмосфере.

**Задача 44.70.** По приведенным данным оценить величину равновесной примеси радиоуглерода в обычном углероде.

**Решение.** Пусть искомая величина примеси равна  $\eta$ . Тогда в одном грамме обычного углерода содержится  $\eta$  граммов радиоуглерода  ${}^{14}\text{C}$ . Исходя из массового числа 14, заключаем, что 14 г радиоуглерода образуют один моль. Значит, в  $\eta$  граммах радиоуглерода содержится  $N = (\eta/14)N_A$  атомов. Скорость распада выражается уравнением  $dN/dt = \lambda N$ , где постоянная распада  $\lambda = \ln 2/T_{1/2} = 3,83 \cdot 10^{-12}$  с. Зная  $dN/dt = 14/60 = 0,233$  с $^{-1}$ , находим:  $\eta = (14/(N_A \lambda))(dN/dt) = 1,4 \cdot 10^{-12}$ , т. е. около  $10^{-10}\%$ . ■

Пока организм живет, его участие в круговороте веществ в природе восполняет убыль радиоуглерода вследствие его распада. В момент смерти организма процесс усвоения прекращается и число атомов радиоуглерода в умершем организме начинает убывать по закону радиоактивного распада:

$$N(t) = N(0)e^{-\lambda t} = N(0)2^{-\frac{t}{T_{1/2}}}.$$

По такому же закону меняется *активность*  $a(t) = dN/dt$  радиоуглерода (*число распадов в единицу времени* — см. следующий раздел). По измеренной активности радиоуглерода в органических образцах можно определить дату их смерти:

$$t = T_{1/2} \frac{\ln(a(0)/a(t))}{\ln 2}. \quad (44.53)$$

Здесь  $a(0)$  — равновесная активность (те самые 14 расп./мин).

В принципе все кажется простым и прозрачным: откапываешь в земле древнюю крепость, сжигаешь кусочек бревна из стены, измеряешь активность образца, и вычисляешь время строительства укрепления. Получили, например,  $a(t) = 13,5$  расп./мин, подставили в формулу (44.53) и нашли возраст крепости:



$t = 5730 \frac{\ln(14/13,5)}{\ln 2} = 300$  лет. Радиоуглеродный метод геохронологии разработан в 1946 г. американцем У. Либби (Нобелевская премия по химии в 1960 г. «за метод использования углерода-14 для определения возраста в археологии, геологии, геофизике и других областях науки»). Метод радиоуглеродных часов прошел проверку и считается одним из самых надежных методов датировки ископаемых органического происхождения. Этому и другим методам датировки исторических артефактов с помощью ядерных реакций распада посвящены монографии<sup>5</sup>, к которым мы отсылаем читателей, интересующихся деталями.

Однако, как и любой научный метод, радиоуглеродный метод имеет свои границы применимости. Несоблюдение «правил игры» приводило иногда к странным результатам: иные живущие организмы, если судить по результатам измерений, должны были давным-давно умереть, а иные ископаемые останки вообще как будто еще не родились на свет. Случаются спекуляции на подобных казусах, имеющие целью посеять недоверие к радиоуглеродному методу и изменить традиционные датировки исторических событий. Поэтому мы кратко обсудим условия, при которых метод дает достоверные результаты.

В основе радиоуглеродного метода лежат следующие допущения.

- Период полураспада  $^{14}\text{C}$  достоверно известен.
- Концентрация  $^{14}\text{C}$  в образце может быть определена с достаточной точностью.
- Изменение соотношения различных изотопов углерода происходит только после смерти организма вследствие радиоактивного распада  $^{14}\text{C}$ .
- Радиоактивный изотоп  $^{14}\text{C}$  быстро и полностью смешивается со стабильными изотопами в глобальных масштабах.
- Отношение количеств  $^{14}\text{C}/^{12}\text{C}$  в глобальном углеродном резервуаре (атмосфере, биосфере, пресной воде, морской воде) остается со временем в основном постоянным.

Два первых предположения особых тревог у физиков не вызывают. Период полураспада  $^{14}\text{C}$  измерен с точностью до 0,5%:  $T_{1/2} = 5730 \pm 40$  лет. Далее надо обеспечить необходимую точность измерений активности образца органического происхождения. Предположим, что при исследовании образца деревянной крепостной стены в задаче 44.70 мы ошиблись в измерениях всего на 1%, т.е. вместо полученных 13,5 расп./мин активность образца на самом деле составляла 13,6 расп./мин. Значит, истинный возраст крепости будет равен  $t = 5730 \frac{\ln(14/13,6)}{\ln 2} = 240$  лет. Ошибка в датировке составила уже 20%.

С погрешностями текущих измерений, как известно, можно бороться путем набора статистики. Оценим, сколько случаев  $n$  распада надо зафиксировать, чтобы активность образца была измерена, скажем, с точностью до 0,5%. Случайный разброс результатов измерений имеет порядок  $\sqrt{n}$ . Относительная точность измерений равна, значит,  $\sqrt{n}/n \sim 0,005$ . Отсюда следует, что надо зафиксировать  $n \sim 40\,000$  распадов. При массе образца 1 г на это потребуется около  $40\,000/13,5 = 3000$  мин = 50 ч. Увеличение же массы образца не всегда возможно.

<sup>5</sup> Г. А. Вагнер. Научные методы датирования в геологии, археологии и истории. Пер. с нем. М.: Техносфера, 2006. М. Walker. Quaternary Dating Methods. J. Wiley & Sons, Ltd., UK, 2005.

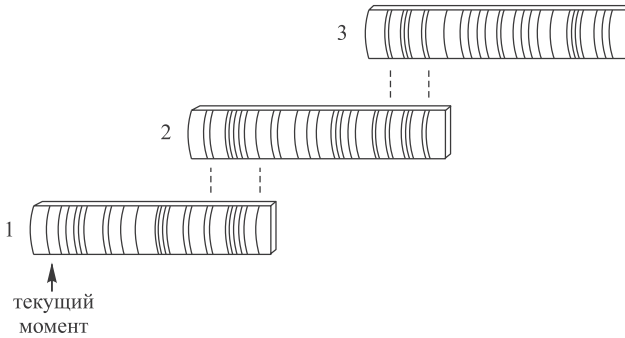
Применение в последние годы масс-спектрометров вместо счетчиков распадов позволяет работать с микроскопическими образцами и получать статистически достоверные результаты.

Третье допущение об изменении соотношения концентраций различных изотопов углерода только из-за радиоактивного распада  $^{14}\text{C}$  после смерти организма может привести к значительным ошибкам в датировке образцов. Химически изотопы углерода неразличимы, они участвуют в одних и тех же химических реакциях. Но из-за разности масс изотопов скорости этих реакций незначительно различаются — происходит обогащение еще живущего образца более легкими изотопами  $^{12}\text{C}$  и  $^{13}\text{C}$ , что приводит к уменьшению относительной концентрации радиоактивного изотопа (эффект *изотопного фракционирования*), а это, в свою очередь, ведет к завышению возраста образца на 80–160 лет (для морских организмов этот эффект может достигать нескольких сотен лет). Его учитывают, измеряя относительную концентрацию стабильных (и более распространенных) изотопов  $^{12}\text{C}$  и  $^{13}\text{C}$ , а затем экстраполируя данные на изотоп  $^{14}\text{C}$ .

Кроме того, происходит перераспределение радиоуглерода между океаном и атмосферой. Океан — это гигантское хранилище  $\text{CO}_2$ , его поверхностные воды быстро (за 5–10 лет) приходят в «углеродное» равновесие с атмосферой. Но время установления равновесия в глубинных водах составляет порядка 2000 лет, и в них концентрация радиоуглерода меньше. При перемешивании вод «старый» углерод поступает в атмосферу, уменьшая относительную концентрацию радиоуглерода, что также приводит к завышению возраста образцов. Влияние океана сказывается на разных оценках возраста образцов из Северного и Южного полушарий — последние оказываются систематически на 30 лет старше. Здесь как раз проявляется влияние океанов, поверхность которых в Южном полушарии больше.

Имеется и техногенное влияние нашей цивилизации. Интенсивное сжигание органического топлива (угля, нефти), начавшееся два века тому назад в результате промышленной революции, также приводит к уменьшению концентрации радиоуглерода, потому что в ископаемом топливе радиоуглерода практически не осталось. Наоборот, после 1950 г. количество радиоуглерода в атмосфере резко возросло из-за испытаний термоядерного оружия. Поэтому радиоуглеродный метод малоприменим для образцов младше 200–300 лет.

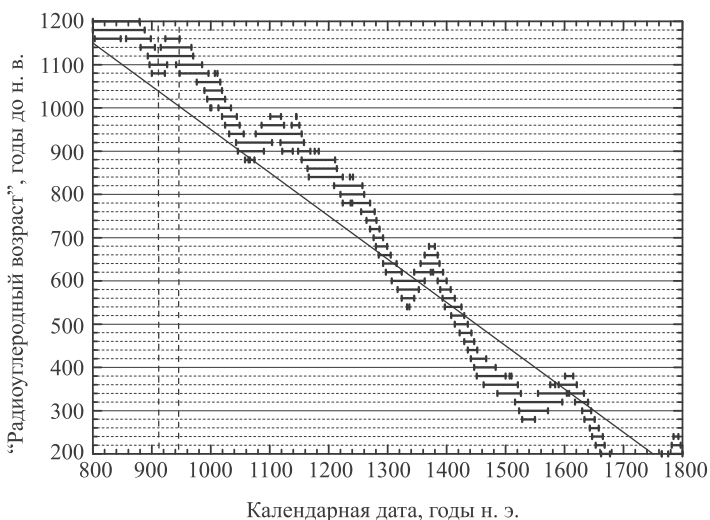
Предположение о постоянстве естественного уровня концентрации  $^{14}\text{C}$  в атмосфере оказалось неверным также из-за вариаций потока космических лучей вследствие изменений уровня солнечной активности и геомагнитного поля, отклоняющего поток заряженных частиц, попадающих в атмосферу. В изначальной версии радиоуглеродного метода начальная равновесная активность  $a(0)$  считалась «почти» не зависящей от времени и местоположения, и в приведенных расчетах ее отождествляли с современным значением равновесной активности радиоуглерода  $A_0$ . Но это «почти» весьма ядовито, поскольку колебания равновесной концентрации (и активности) радиоуглерода могут достигать нескольких процентов. Мы уже знаем, что отклонение в активности образца на 1% приводит к ошибке в датировке на 20%. А что получится, если по каким-то причинам равновесная концентрация радиоуглерода во время и в месте строительства древней крепости (см. задачу 44.70) отличалась от нынешней на 3% и составляла  $a(0) = 14,44$  расп./мин? Мы этого не знали, использовали



**Рис. 44.7.** Метод дендрохронологии: сравнение последовательности годичных колец и образцов спилов деревьев разного возраста позволило протянуть непрерывную нить от наших дней в глубь времен на много тысяч лет назад

вместо  $a(0)$  величину  $A_0$ , и определили возраст крепости в 300 лет, а на самом деле он должен бы получиться равным  $t = 5730 \frac{\ln(14,44/13,5)}{\ln 2} = 556$  лет. Отклонение начальной активности  $a(0)$  радиоуглерода от современной равновесной  $a_0 = 14$  распадов/мин всего на три процента привело к ошибке в датировке почти в два раза! В результате мы рискуем приписать Петру I то, что построено при Золотой Орде.

Стало ясно, что перед применением радиоуглеродного метода необходимо заняться *калибровкой* — установить начальные концентрации радиоуглерода в атмосфере для различных годов. И эта принципиальная проблема оказалась разрешимой. Дело в том, что информация о начальной концентрации радиоуглерода накапливается в годичных кольцах деревьев. Все годичные кольца, кроме внешнего, в обмене уже не участвуют: они «омертвели», сохранив информацию о концентрации радиоуглерода в тот год, когда они образовывались. Попавший в кольцо радиоуглерод распадается далее по закону радиоактивного распада. Взяв образец древесины из кольца, чья датировка надежно установлена, и измерив его активность в настоящее время, можно найти концентрацию радиоуглерода в момент образования данного кольца. В западных штатах США растет остистая сосна (*Pinus longaeva*) — самое долгоживущее дерево на Земле (более 4000 лет). При этом сосна растет так медленно, что редко превышает 10 м в высоту. В Национальном парке *White Mountains* («Белогорье» в русском переводе) на границе Калифорнии и Невады до сих пор растет дерево, получившее имя собственное «Мафусаил». Приблизительная оценка времени прорастания семени, из которого оно выросло, — 2831 г. до н. э. (т. е. по состоянию на 2016 г. примерный возраст дерева составляет 4845 лет). Подумать только, что в наше время можно потрогать живое дерево, которое начало расти во времена возникновения египетского государства и правления у шумеров Гильгамеша (правда, местоположение дерева держится в секрете во избежание актов вандализма). Исследуя эти долгоживущие деревья, а также немецкие и ирландские дубы, ученые проследили образование годичных колец на 10–12 тыс. лет назад и сумели установить начальные концентрации радиоуглерода для этого периода.



**Рис. 44.8.** Калибровочные кривые для периода 800–1700 гг. н.э. По оси ординат отложен «радиоуглеродный возраст» образца; горизонтальные отрезки с вероятностью 95,5% указывают нижнюю и верхнюю границы соответствующей ему календарной даты. Вертикальные штриховые линии отмечают даты смерти князей Олега (†912) и Игоря (†945). Калибровочные кривые превратились бы в диагональную линию  $y = 1950 - x$ , если бы «радиоуглеродный возраст» точно соответствовал календарному. Расчеты калибровочных кривых выполнены на программе Calib 7.0.4 (M. Stuiver, P.J. Reimer, R. Reimer), см. сайт <http://calib.qub.ac.uk/calib/>

Принцип такой *дендрохронологии* объяснен на рис. 44.7. На нем показаны три образца, выпиленные из стволов деревьев (слева — внешняя часть, ближняя к коре, справа — сердцевина), на которых имеются годовичные кольца. В зависимости от климатических условий кольца могут быть шире или уже, образуя своего рода отпечатки пальцев того периода, когда жило дерево. Образец 1 взят из ныне живущего дерева, так что датировка внешнего кольца соответствует текущему моменту. Образец 2 взят из стоящего, но давно умершего дерева. В нем обнаружена та же последовательность колец, что и в первом образце, что позволяет отождествить соответствующие годы (показано пунктирными линиями). Но если в образце 1 эта последовательность колец соответствовала молодости дерева, то в образце 2 она образовалась уже под конец жизни. Поэтому образец 2 позволяет установить более раннюю последовательность годовых колец. Образец 3 еще старше образца 2, и отождествление последовательности колец этих двух деревьев позволяет еще дальше «продвинуться» в прошлое. Таким путем были установлены датировки годовичных колец за последние 12 000 лет.

Итак, определив значения равновесной концентрации радиоуглерода для разных моментов времени, ученые обнаружили, что они действительно меняются, иногда даже в пределах 5%. Итогом подобных исследований стали калибровочные кривые, примерный вид которых показан на рис. 44.8. Тут необходимы пояснения. По вертикальной оси отложен «радиоуглеродный возраст» исследуемого образца в годах до н. в. — «до настоящего времени». Точкой отсчета для

«настоящего времени» считают 1 января 1950 г., и возраст образца условились отсчитывать от этой даты. По горизонтальной оси дата может откладываться также в годах до н. в., или, как на рис. 44.8, в обычных календарных годах. Но радиоуглеродный метод не сопоставляет однозначно «радиоуглеродный возраст» с календарной датой, а дает лишь интервал возможных значений последней (на рис. 44.8 эти интервалы показаны горизонтальными отрезками, соответствующими удвоенной среднеквадратической погрешности  $2\sigma$ , т. е. вероятности 95,45%).

Пусть исследуется образец, для которого радиоуглеродный метод дает возраст 727 лет до н. в., что относится к 1223 г. н. э. Но не надо спешить радоваться, что найден артефакт времен битвы на Калке. Расчет по программе Calib 7.0.4, с помощью которой построена диаграмма на рис. 44.8, дает для этого значения «радиоуглеродного возраста» интервал календарных дат от 1267 до 1284 г.

Следует обратить внимание на то, что волнообразный характер кривых приводит к тому, что для образца могут получиться несколько возможных дат. Так, если «радиоуглеродный возраст» образца оказывается равным 900 годам, что будто бы соответствует концу правления киевского князя Ярослава Мудрого († 1054), то калибровочные кривые указывают на четыре (!) возможных интервала календарных дат, из которых наиболее вероятны 1046–1090 гг. и 1148–1169 гг., но возможны также 1121–1139 гг. и 1176–1183 гг.

Прежде чем проводить измерения с калибровочными кривыми, консультируются, чтобы понять, может ли радиоуглеродный метод ответить на поставленный вопрос. Скажем, производятся раскопки кургана и имеются подозрения, что он является захоронением либо князя Игоря († 945), либо его предшественника князя Олега († 912). Разрешит ли эту дилемму исследование радиоуглеродным методом найденных там образцов? Проводя на рисунке 44.8 вертикальную прямую, проходящую через точку с координатами  $x = 912$ , видим, что она пересекает отрезки, соответствующие возрасту образца в интервале 1080–1120 лет до н. в., так что календарная дата лежит в интервале 895–935 гг. н. э. Аналогичное построение для точки  $x = 945$  дает возраст образца в пределах 1080–1160 лет до н. в., что соответствует календарным датам 855–935 гг. н. э. Указанные интервалы сильно пересекаются, так что для ответа на данный вопрос радиоуглеродный метод бесполезен и не стоит тратить время и деньги (это — недешевое удовольствие) для проведения измерений.

Следует также иметь в виду, что радиоуглеродный метод в принципе работает для образцов не старше 50–60 тыс. лет. Дело в том, что при нормальных условиях измерениям доступны количества углерода-14, оставшиеся после примерно восьми периодов полураспада (45 тыс. лет). Ухищрения разного рода позволяют увеличить этот интервал до 50–60 тыс. лет — это предел «дальнейности» радиоуглеродного метода. В более древних образцах радиоуглерода осталось так мало, что измерения его количества невозможны. В целом метод очень надежно работает для последних тысячелетий. Исследование образцов (в частности, кораллов) с помощью радиоактивных изотопов урана и продуктов его распада позволило продлить калибровочные кривые за дендрочасток.

Из нашедших применений радиоуглеродного метода упомянем датировки найденных на Мертвом море свитков (100 г. до н. э. — 100 г. н. э.) и замороженного тела, найденного в 1991 г. в Италии (3300–3100 гг. до н. э.). Один из наиболее

скандальных результатов, полученных радиоуглеродным методом, связан с датировкой туринской плащаницы, по преданию — погребальной одежды Иисуса Христа. Исторические упоминания об этой реликвии в литературе относятся к середине XIV в. В 1980-х гг. архиепископ Турина дал разрешение на исследование маленьких образцов ткани плащаницы. Работы проводились в трех лабораториях — в США, Англии и Швейцарии. Результаты всех групп согласуются друг с другом: с вероятностью 95% плащаница датируется периодом 1262–1394 гг. Высказанные соображения о возможности загрязнения плащаницы более молодым радиоуглеродом во время пожаров были проанализированы в 1996 г. и отклонены как необоснованные. Но вероятность возможного загрязнения образца — одна из трудностей радиоуглеродного метода, поэтому разработаны разные методы очистки. К загрязнению особенно чувствительны образцы с малым содержанием радиоуглерода. Так, для образца возрастом 40 тыс. лет загрязнение современным углеродом всего в 1% уменьшит его радиоуглеродный возраст на 7000 лет.

В целом сочетание различных радиохимических методов датировки вместе с данными геологии позволяет установить надежную хронологическую шкалу для всего четвертичного периода, начавшегося 2,6 млн лет назад, продолжающегося до сих пор и ознаменованного самым важным для нас событием в истории Земли — появлением рода *Ното*, одним из видов которого является человек.

Для датировок неорганических образцов используют другие радиоактивные изотопы. Известен, например, калий-аргоновый метод. Калий — первичный породообразующий элемент, он присутствует на Земле со времени ее образования. Равновесная концентрация радиоактивного изотопа калий-40 в природном калии составляет 0,0117%. Уже упоминалось, что у этого изотопа есть два канала распада:  $\beta$ -распад  ${}^{40}_{19}\text{K} \rightarrow {}^{40}_{20}\text{Ca} + e^- + \bar{\nu}_e$  и захват орбитального электрона  ${}^{40}_{19}\text{K} + e^- \rightarrow {}^{40}_{18}\text{Ar} + \nu_e + \gamma$ . Большой период полураспада по «аргоновому» каналу (11,7 млрд лет, см. задачу 44.68) делает возможным его использование для датировки очень древних образцов. Аргон-40 — стабильный инертный газ, он легко высвобождается при нагревании, и по оставшемуся его количеству в образце можно судить о времени, когда такое нагревание случилось в последний раз.

Разновидность этого метода — аргон-аргоновая датировка. В нем измерение количества калия-40 заменяют на измерение соотношения количеств аргона-40 и аргона-39. Период полураспада аргона-39 — 269 лет, так что в старых образцах его быть не может. Образец подвергают в реакторе облучению быстрыми нейтронами, изотоп  ${}^{39}\text{Ar}$  образуется в реакции  ${}^{39}\text{K}(n, p){}^{39}\text{Ar}$ , и по его количеству можно судить о количестве  ${}^{40}\text{K}$ . Аргон-аргоновый метод сложнее, но у него есть свои преимущества. Именно таким образом в 1997 г. были датированы образцы вулканической пемзы из раскопок Помпей. Полученный результат для возраста образцов (1925 лет  $\pm$  94 года) привел к дате  $72 \pm 94$  г. н. э. извержения Везувия, погубившего этот древнеримский город, что превосходно согласуется с традиционной датировкой (79 г. н. э.) по историческим источникам. Этим же методом были датированы породы в Эфиопии, содержащие останки первобытных людей, которых по анатомическому строению относят уже к нашим прямым предкам — виду *Homo Sapiens*: 160 000–154 000 г. до н. э.

Для датировки используются и другие медленные распады. Например, свинец-свинцовый метод, основанный на сравнении концентраций трех изотопов свинца — одного стабильного ( $^{204}\text{Pb}$ ) и двух радиоактивных ( $^{206}\text{Pb}$ , образующегося при распаде  $^{238}\text{U}$ , и  $^{207}\text{Pb}$ , возникающего при распаде  $^{235}\text{U}$ ) — дал оценку  $4,54 \pm 0,05$  млрд лет для возраста Земли.

**Задача 44.71.** Анализ на масс-спектрометре образцов породы, доставленной с Луны, показал, что отношение количеств атомов  $^{40}\text{Ar}$  и  $^{40}\text{K}$  равно  $\eta = 10,3$ . Каков возраст лунной породы?

**Решение.** Радиоактивный изотоп калия попал в породу в момент ее образования, и пусть его начальное количество составляло  $N_{\text{K},0}$ . К моменту времени  $t$  число атомов калия составит  $N_{\text{K}} = N_{\text{K},0}2^{-t/T_{1/2}}$ . Число же атомов аргона, который целиком образовался из распавшегося калия, равно  $N_{\text{Ar}} = N_{\text{K},0} - N_{\text{K}}$ . Отношение количеств аргона и калия равно  $\eta = N_{\text{Ar}}/N_{\text{K}} = N_{\text{K},0}/N_{\text{K}} - 1 = 2^{t/T_{1/2}} - 1$ , откуда находим:

$$t = T_{1/2} \frac{\ln(1 + \eta)}{\ln 2} = 4,37 \cdot 10^9 \text{ лет.} \quad \blacksquare$$

## 44.6 Элементы дозиметрии

В предыдущем разделе мы уже использовали понятие активности радиоактивного образца в связи с определением концентрации радиоактивных ядер в методах датировки исторических артефактов и горных пород. Для дозиметрии это понятие также является ключевым, поэтому мы напомним, что

♦ **активностью радиоактивного образца называется число распадов, происходящих в нем в единицу времени.** ♦

Рассмотрим более подробно закон изменения активности во времени. Если за время  $dt$  распадается  $dN_{\text{расп}}$  ядер, то активность определяется соотношением  $dN_{\text{расп}}/dt$ . Поскольку убыль ядер радиоактивного вещества равна по величине и обратна по знаку числу распавшихся ядер ( $dN_{\text{расп}} = -dN$ ), активность  $a$  дается выражением:

$$a = -\frac{dN}{dt} = \lambda N = \frac{\ln 2}{T_{1/2}} N. \quad (44.54)$$

Таким образом, активность радиоактивного вещества пропорциональна количеству радиоактивных ядер и обратно пропорциональна периоду полураспада. Соотношение (44.44) означает, что в веществе устанавливается такая концентрация ядер, при которой равны активности всех радиоактивных изотопов в препарате.

♦ **Активность в расчете на единицу массы источника называется удельной активностью.** ♦

В СИ за единицу активности принимают один распад в секунду. Эта единица носит название *беккерель* (Бк). В качестве внесистемной единицы активности используется единица, называемая *кюри*, соответствующая активности образца, в котором происходит  $3,7 \cdot 10^{10}$  распадов в секунду, т. е.  $1 \text{ Ки} = 3,7 \cdot 10^{10} \text{ Бк} = 37 \text{ ГБк}$ . Происхождение этой единицы следующее. Ранее мы рассматривали пример распада радия и образовавшегося радона. Если в закрытый сосуд поместить радий, то сначала количество радона будет возрастать, а затем установится

равновесие между вновь возникающим радоном и распадающимися ядрами радона, т.е. число ежесекундно совершающихся актов распада будет оставаться практически постоянным, если не учитывать изменение массы «материнского» радия, которое происходит весьма медленно вследствие большого периода полураспада (1622 года). Единица активности кюри (1 Ки) представляет собой радиоактивность радона, находящегося в равновесии с одним граммом радия.

**Задача 44.72.** Изотоп золота  $^{198}\text{Au}$ , используемый для терапии рака, имеет период полураспада  $T_{1/2} = 2,7$  сут. Вычислить массу препарата, необходимую для получения активности  $a = 250$  Ки.

**Решение.** Исходя из определения активности, сразу находим из (44.54), что необходимое количество атомов равно  $N = aT_{1/2}/\ln 2$ . Период полураспада, выраженный в секундах, равен  $T_{1/2} = 2,7 \times 86\,400 = 2,33 \cdot 10^5$  с. Активность  $a = 250$  Ки  $= 9,25 \cdot 10^{12}$  Бк. Подставляя эти численные данные, находим:  $N = 3,11 \cdot 10^{18}$ . В расчете подобного рода разница в массах протона и нейтрона и дефект массы ядра несуществен, так что массу атома золота находим как  $m_{\text{Au}} = (Z + N)m_p = 198 \times 1,67 \cdot 10^{-27}$  кг  $= 3,31 \cdot 10^{-25}$  кг. Искомая масса препарата тогда равна  $m = Nm_{\text{Au}} = 1,03 \cdot 10^{-6}$  кг  $= 1,03$  мг. ■

Образующиеся при радиоактивных распадах заряженные частицы и  $\gamma$ -кванты, пролетая через вещество, постепенно теряют свою энергию, затрачивая ее на ионизацию молекул вещества (именно отсюда происходит термин «ионизирующие излучения»), и, в конце концов, останавливаются. Например, на образование в воздухе одной пары ионов противоположных знаков тратится в среднем энергия, равная 34 эВ (из этой энергии на ионизацию молекул идет несколько меньше половины, а остальная энергия излучения тратится на возбуждение молекул и их упругие столкновения).

Поэтому при исследовании *ионизирующих излучений* особое значение имеет величина энергии, поглощенной единицей массы данного вещества. **Поглощенной дозой  $D$**  излучения называется энергия, поглощенная единицей массы облучаемого объекта. Ее размерность  $[D] = [E/m] = \text{L}^2\text{T}^{-2} = \text{Дж/кг}$ . В СИ эта единица получила название *грей* (Гр). Часто используется также внесистемная единица *рад*: 1 рад  $= 0,01$  Дж/кг, т.е. 1 Гр  $= 100$  рад. Рад — крупная единица, и потому поглощенные дозы обычно выражают в долях рада — сотых (сантирад), тысячных (миллирад) и миллионных (микрорад).

Человеческий организм поглощает энергию ионизирующих излучений, причем от количества поглощенной энергии зависит степень лучевых поражений. Но при облучении живых организмов возникающие биологические эффекты зависят не только от поглощенной дозы, но и от вида излучения. Степень опасности облучения для живого организма выражается в понятии **эквивалентной дозы**. Чтобы получить значение эквивалентной дозы, надо величину поглощенной дозы умножить на **коэффициент качества излучения  $Q$** . Биологическое влияние излучения принято сравнивать с эффектами от рентгеновского излучения или гамма-излучения, которое берется за эталон. Для него по определению полагают  $Q = 1$ . Единицей измерения эквивалентной дозы в СИ служит **зиверт** (Зв): 1 Зв  $= 1$  Гр  $\cdot Q$ . Используют также внесистемную единицу **бэр** (биологический эквивалент рентгена): 1 бэр  $= 1$  рад  $\cdot Q$ , т.е. 1 Зв  $= 100$  бэр. Таким образом, если организм получил дозу в 1 Зв в виде рентгеновских лучей, то полученная эквивалентная доза составит 100 бэр. Однако при той же поглощенной дозе при



облучении нейтронами, для которых  $Q = 3$ , эквивалентная доза составляет уже 300 бэр, т. е. в три раза опаснее.

**Задача 44.73.** Типичная эквивалентная доза при рентгенокопии грудной клетки составляет  $D_{\text{экв}} = 25$  мбэр. Предполагая, что масса человека  $m = 80$  кг и облучается примерно половина тела, найти поглощенную энергию  $W$ .

**Решение.** В СИ эквивалентная доза равна  $D_{\text{экв}} = 25 \cdot 10^{-3}/100 = 2,5 \cdot 10^{-4}$  Зв. Поглощенная доза составит при этом  $D = D_{\text{экв}}/Q = 2,5 \cdot 10^{-4}$  Гр. Умножая на массу  $m/2$  облученной части тела, находим поглощенную энергию:  $W = mD/2 = 0,01$  Дж. ■

Коэффициент качества излучения отражает степень риска (возникновения рака, лейкемии и т. п.) при хроническом облучении малыми дозами. Гамма-излучение сопровождается многими ядерными реакциями и распадами. Энергия  $\gamma$ -квантов может принимать значения от кэВ до МэВ. Гамма-излучение отличается высокой проникающей способностью, и для защиты от него эффективно применяют материалы с большим удельным весом: свинец, бетон и т. п. Для  $\beta$ -лучей коэффициент качества излучения примерно такой же:  $Q \approx 1$ . Этот вид излучения — поток электронов и позитронов, т. е. частиц, обладающих электрическим зарядом и, следовательно, высокой ионизирующей способностью. Они могут проходить слой воды толщиной 1–2 см. Для защиты от  $\beta$ -излучения обычно достаточно листа алюминия толщиной в несколько миллиметров. Но попадание источников  $\beta$ -излучения в организм с пищей, водой и воздухом может привести к опасному внутреннему облучению организма. Альфа-излучение — поток ядер (ионов) гелия, возникающих при радиоактивном распаде элементов с большими атомными номерами. Альфа-частицы в 7300 раз тяжелее электронов и несут по сравнению с ними двойной электрический заряд. Поэтому их ионизирующая способность чрезвычайно велика, как и у других тяжелых ионов, но они легко задерживаются даже листом бумаги. Пробег  $\alpha$ -частиц в воздухе не превышает 11 см. От внешнего  $\alpha$ -излучения легко защититься, но оно чрезвычайно опасно при внутреннем облучении. Для  $\alpha$ -излучения и облучения тяжелыми ионами  $Q \approx 20$ . При оценке радиоактивной обстановки существенную роль может играть нейтронное излучение. Нейтроны, например, вылетают из ядер при реакциях деления урана и плутония. Сами они электрически нейтральны и имеют высокую проникающую способность. При взаимодействии нейтронов с веществом возникают заряженные частицы, ионизирующие среду. Нейтронное излучение способно превращать атомы стабильных элементов в их радиоактивные изотопы, что повышает опасность нейтронного излучения. От него хорошо защищают материалы, содержащие водород: парафин, вода и т. п. Для тепловых (медленных) нейтронов  $Q \approx 5$ , для быстрых —  $Q \approx 10$ –20 в зависимости от энергии нейтронов.

Органы человека имеют разную чувствительность к радиации. При той же эквивалентной дозе возникновение рака при облучении легких на порядок более вероятно, чем при облучении кожи. Особенно опасно облучение половых органов, что может привести к тяжелым генетическим последствиям. Чувствительность отдельных органов к радиации выражается в виде *коэффициентов радиационного риска*, величина которых установлена эмпирически (табл. 44.2) Их сумма равна единице, и с их помощью вычисляется *эффективная доза*  $E$ , равная сумме произведений эквивалентных доз, полученных отдельными органами, на

**Таблица 44.2.** Коэффициенты радиационного риска для различных органов (по данным Википедии)

Органы, ткани	Коэффициент	Органы, ткани	Коэффициент
Половые железы	0,20	Печень	0,05
Красный костный мозг	0,12	Пищевод	0,05
Толстый кишечник	0,12	Щитовидная железа	0,05
Желудок	0,12	Кожа	0,01
Легкие	0,12	Костные поверхности	0,01
Мочевой пузырь	0,05	Головной мозг	0,05
		Остальные ткани	0,05

соответствующие коэффициенты риска. Эффективная доза отражает суммарный эффект облучения для всего организма. Так как коэффициенты радиационного риска безразмерны, единицы измерения эффективной дозы и эквивалентной дозы совпадают.

Поглощенная и эквивалентная дозы характеризуют влияние радиоактивной среды на облучаемые объекты. Для характеристики радиационной обстановки на местности и в помещениях используют понятие *экспозиционной дозы*: это количество рентгеновского или  $\gamma$ -излучения, выраженное в единицах заряда одного знака, порожденного им в единице массы сухого воздуха в данной точке пространства. Единицей экспозиционной дозы в СИ является Кл/кг. На практике часто используют внесистемную единицу — рентген (Р):  $1 \text{ Р} = 2,58 \cdot 10^{-4} \text{ Кл/кг}$ . Так как на образование в воздухе пары ионов, т. е. заряда  $e$  (считаются заряды ионов одного знака) тратится 34 эВ, энергия, затрачиваемая на образование единицы заряда, равна  $w = 34 \text{ Дж/Кл}$ . Теперь экспозиционную дозу можно связать с поглощенной (воздухом же) дозой:  $D_{\text{возд}} = 1 \text{ Р} \times w = 2,58 \cdot 10^{-4} \times 34 = 8,77 \text{ мГр} = 0,877 \text{ рад}$ . Однако это не означает, что такую же дозу получает находящийся там объект. По величине экспозиционной дозы можно рассчитать поглощенную любым объектом дозу, зная характеристики излучения, состав, форму и др. свойства объекта. Поэтому нельзя говорить даже о примерной эквивалентности одного рентгена, рада или бэра: это единицы введены для разных характеристик излучения и оказываемого им влияния.

При прочих равных условиях доза ионизирующего излучения тем больше, чем больше время облучения. Поэтому используют также понятие *мощности дозы*: это соответствующая (экспозиционная, поглощенная или эквивалентная) доза, отнесенная к единице времени:  $1 \text{ Р/ч}$  и т. п.

Предельно допустимая эквивалентная доза для профессиональных работников при облучении всего тела установлена в 0,1 бэр (1 мЗв) в неделю. Считается, что такая доза при систематическом воздействии в течение неограниченно продолжительного времени не вызывает у работающих каких-либо патологических изменений или заболеваний. При шести часах работы пять дней в неделю это соответствует мощности эквивалентной дозы  $H_{\text{доп}} = 30 \text{ мкЗв/ч}$ .

При облучении  $\gamma$ -квантами с энергией  $E_\gamma$  мощность предельно допустимой эквивалентной дозы можно связать с предельно допустимой плотностью потока

фотонов  $J_{\text{доп}}$ :

$$H_{\text{доп}} = J_{\text{доп}} E_{\gamma} \mu_m / Q. \quad (44.55)$$

Здесь  $Q$  — коэффициент качества излучения ( $Q = 1$  для фотонов), а  $\mu$  — **массовый коэффициент поглощения**, связанный с линейным коэффициентом поглощения  $\kappa$  в законе Бугера (см. разд. 31.3) соотношением  $\mu = \kappa / \rho$ , где  $\rho$  — плотность вещества мишени. Так, для  $\gamma$ -квантов с энергией порядка 1–2 МэВ линейный коэффициент поглощения воды равен  $\kappa_{\text{H}_2\text{O}} = 0,06 \text{ см}^{-1}$ , а  $\rho_{\text{H}_2\text{O}} = 1 \text{ г/см}^3$ , откуда  $\mu_{\text{H}_2\text{O}} = 0,006 \text{ м}^2/\text{кг}$ . Поскольку человек состоит в основном из воды, мы примем это значение за оценку массового коэффициента поглощения в биологической ткани. Отсюда можно получить величину предельно допустимого потока фотонов:

$$J_{\text{доп}} = 5 \cdot 10^{-3} \left( \frac{\text{Дж}}{E_{\gamma}} \right) \text{ м}^{-2} \text{ч}^{-1} = 1,4 \cdot 10^{-6} \left( \frac{\text{Дж}}{E_{\gamma}} \right) \text{ м}^{-2} \text{с}^{-1}. \quad (44.56)$$

**Задача 44.74.** При лечении некоторых видов рака используют излучение изотопа кобальта  $^{60}\text{Co}$ , при  $\beta$ -распаде которого излучаются фотоны с энергией  $E_{\gamma} = 1,33 \text{ МэВ}$ . Пусть радиоактивный источник активностью  $a = 10 \text{ мКи}$  помещен в сферический свинцовый контейнер. Какова должна быть толщина  $l$  стенок контейнера, чтобы для оператора, работающего на расстоянии  $R = 1 \text{ м}$  от источника радиации, соблюдались предельно допустимые дозы? Для излучения указанной энергии коэффициент линейного поглощения в свинце равен  $\kappa_{\text{Pb}} = 0,64 \text{ см}^{-1}$ .

**Решение.** Если в каждом распаде излучается один фотон, то число фотонов, испущенных радиоактивным препаратом в единицу времени, равно активности  $a$  препарата. Тогда плотность потока фотонов на расстоянии  $R$  от источника равна  $J = a / (4\pi R^2)$ . При прохождении слоя свинца толщиной  $l$  в соответствии с законом Бугера поток ослабляется, и его значение на выходе должно быть не больше предельно допустимой нормы:

$$J e^{-\kappa l} \leq J_{\text{доп}}. \quad (44.57)$$

Отсюда получаем оценку для толщины стенок контейнера:

$$l \geq \frac{1}{\kappa} \ln \frac{J}{J_{\text{доп}}}. \quad (44.58)$$

Найдем численное значение потока  $\gamma$ -квантов от препарата:

$$J = \frac{a}{4\pi R^2} = \frac{10^{-2} \times 3,7 \cdot 10^{10}}{4\pi} = 2,9 \cdot 10^7 \frac{1}{\text{м}^2 \text{с}}. \quad (44.59)$$

Предельно допустимую плотность потока  $\gamma$ -квантов находим, подставляя их энергию  $E_{\gamma} = 1,33 \text{ МэВ}$  в (44.56) и учитывая соотношение  $1 \text{ МэВ} = 1,6 \cdot 10^{-13} \text{ Дж}$ :

$$J_{\text{доп}} = \frac{1,4 \cdot 10^{-6}}{1,33 \times 1,6 \cdot 10^{-13}} = 6,6 \cdot 10^6 \frac{1}{\text{м}^2 \text{с}}. \quad (44.60)$$

Подставляем теперь результаты (44.59) и (44.60) в (44.58):

$$l = \frac{1}{0,64} \ln \frac{29}{6,6} = 2,3 \text{ см}. \quad \blacksquare \quad (44.61)$$

Средняя мощность экспозиционной дозы от всех естественных источников ионизирующего излучения составляет около  $200 \text{ мР/год}$  (т.е. примерно равна

Таблица 44.3. Природные источники ионизирующего излучения

Источники	Мощность дозы, мбэр/год	Доля источника, %
Космическое излучение на уровне моря	30	26,8–15,6
Поверхность Земли (грунт, вода)	50–130	44,6–67,7
Радиоактивные элементы в теле человека	30	26,8–15,6
Прочие источники	2	1,8–1,1
Итого	112–192	100

Таблица 44.4. Средние годовые дозы от искусственных источников излучения

Источники	Мощность дозы, мбэр/год	Доля от фона, %
Медицинская диагностика	100–150	50–75
Строительные материалы (пробывание в зданиях)	150	75
Полеты в самолете (10 000 км на высоте 12 км)	2,5–5	1–2,5
Телевизор (4 ч/сут.)	1	0,5
ТЭЦ (1 ГВт) на угле на расстоянии до 20 км	0,6–6	0,3–3
АЭС (1 ГВт) на расстоянии до 20 км	0,01–0,1	≤ 0,1
Радиоактивные осадки от испытаний ядерного оружия	2,5	1
Прочие источники	40	20
Итого:	300–350	150–175

20 мкР/ч). Примерно такие же значения можно слышать в сообщениях по ТВ о радиационной обстановке в городах. В разных областях Земли мощность экспозиционной дозы колеблется от 50 до 1000 мР/год и более (6–110 мкР/ч). Экспозиционная доза 200 мР за год примерно соответствует эквивалентной дозе 200 мбэр за год. В табл. 44.3 показано распределение естественной экспозиционной дозы по источникам происхождения.

В целом приведенные дозы представляют собой естественную среду обитания человека, а указанные факторы облучения действуют уже миллионы лет, создавая природный фон ионизирующего излучения. Конечно, деятельность людей вносит свои коррективы. В естественных горных породах, в почве и строительных материалах содержится какое-то количество радия, при распаде которого возникает радиоактивный газ радон. Плохая вентиляция может увеличить дозу облучения, обусловленную вдыханием радона. Доза от воздействия космических лучей зависит от высоты над уровнем моря. Дополнительному облучению подвергаются, например, пассажиры авиалайнеров. Имеются и другие искусственные источники облучения, приведенные в таблице 44.4. За естественный фон здесь принято значение 200 мбэр/год. Видно, что деятельность человека примерно утраивает эквивалентную дозу. Основным источником дополнительного облучения являются наше жилье и медицинские приборы. При флюорографии пациент

получает около 370 мбэр, при рентгенографии зуба — 3 мбэр, при рентгеноскопии легких — 2–8 бэр!

При эксплуатации АЭС нормативными документами установлены пределы для облучения персонала и населения: соответственно 5 и 0,5 бэр/год. За 70 лет жизни от естественного фона набирается полная доза около 15 бэр. При установлении указанного норматива для населения за основу был взят предельный уровень 35 бэр за жизнь от всех источников излучения. Такая доза считается потенциально неопасной.

При выполнении аварийных работ максимально накопленная доза облучения не должна превышать 25 бэр. Кратковременные незначительные изменения состава крови наблюдались при разовых дозах 25–75 бэр. Летальной дозой считается доза проникающего ионизирующего излучения, приводящая с 50%-й вероятностью к летальному исходу в течение 30 дней в отсутствие медицинской помощи. Она составляет 250–300 бэр. Предельная доза внешнего облучения может варьироваться в зависимости от условий облучения и свойств излучения.

В заключение приведем любопытный расчет, появившийся в научной литературе в связи с некоторыми теориями, предсказывавшими, что протон может оказаться нестабильным. Потребовались экспериментальные оценки на время жизни протона, и одна из них может быть легко получена на основании только что приведенных данных. Для грубой оценки примем, что человек массой  $m = 80$  кг наполовину состоит из протонов. Их число  $N_p$  можно найти, разделив  $m/2$  на массу протона:  $N_p = m/2m_p = 40/(1,67 \cdot 10^{-27}) = 2,4 \cdot 10^{28}$ . При распаде протона выделяется его энергия покоя  $m_p c^2 = 1$  ГэВ =  $1,6 \cdot 10^{-10}$  Дж. Если среднее время жизни протона составляет  $\tau$  (в годах), то вероятность распада протона за время жизни человека (70 лет) равна  $W_p \sim 70/\tau$ . Тогда за 70 лет в теле распадется  $n = W_p N_p = 1,7 \cdot 10^{30}/\tau$  протонов, и поглощенная доза составит  $(m_p c^2 n)/m = 3,4 \cdot 10^{18}/\tau$  Гр. При распаде протона должны образовываться легкие частицы ( $Q = 1$ ), т. е. эквивалентная доза будет равна  $D = 3,4 \cdot 10^{20}/\tau$  бэр. Эта величина не должна превышать биологически опасного предельного значения 35 бэр (ведь человечество не вымерло от распада протонов, даже если он и происходит). Неравенство  $D \leq 35$  бэр дает оценку  $\tau \geq (3,4 \cdot 10^{20}/35) \approx 10^{19}$  лет. Конечно, физические эксперименты позволили поднять эту оценку на много порядков, но уже этот простой расчет дает невероятно большое значение для времени жизни протона: наша Вселенная существует «всего»  $13,7 \cdot 10^9$  лет. Так что на нашу обычную жизнь нестабильность протона, если она и имеет место, никакого влияния оказать не может.

## 44.7 Атомная и термоядерная энергетика

### Деление ядер

Из следствий модели жидкой капли остановимся на критерии стабильности относительно деформации ядра. Представим себе, что первоначально сферическое ядро деформировалось и вытянулось в эллипсоид с полуосями  $a = r_\alpha(1 + \epsilon)$ ,  $b = r_\alpha/\sqrt{1 + \epsilon}$  (при  $\epsilon = 0$  величины полуосей совпадают, и мы имеем рассмотренный случай сферического ядра). Приведенные выражения для полуосей получаются в предположении, что объем ядра при деформации не изменился. В выражении для массы  $M_\epsilon(Z, A)$  деформированного ядра изменятся два слагаемых. Во-первых, деформация увеличивает площадь поверхности ядра: она становится равной  $4\pi r_\alpha^2(1 + 2\epsilon^2/5)$ . Соответствующее слагаемое в формуле (44.20) примет теперь вид  $0,0180A^{2/3}(1 + 2\epsilon^2/5)$ . Это приводит к изменению массы ядра

на величину

$$\Delta_S M = 0,0072 A^{2/3} \epsilon^2.$$

Во-вторых, меняется энергия кулоновского отталкивания протонов. Она должна уменьшиться, так как среднее расстояние между протонами становится больше. Электростатическая энергия эллипсоида имеет вид

$$E_C = \frac{3}{5} \cdot \frac{e^2 Z^2}{4\pi\epsilon_0 r_{я}} \left(1 - \frac{\epsilon^2}{5}\right),$$

так что кулоновское слагаемое в формуле (44.20) примет вид

$$0,000713 \cdot (Z^2/A^{1/3}) \cdot (1 - \epsilon^2/5).$$

Это приводит к изменению массы ядра на величину

$$\Delta_C M = -0,0001426 (Z^2/A^{1/3}) \epsilon^2.$$

Теперь находим полное изменение массы ядра при деформации:

$$\begin{aligned} \Delta_\epsilon M &= M(Z, A) - M_\epsilon(Z, A) = -\Delta_S M - \Delta_C M = \\ &= \epsilon^2 \left( -0,0072 A^{2/3} + 0,0001426 \frac{Z^2}{A^{1/3}} \right). \end{aligned} \quad (44.62)$$

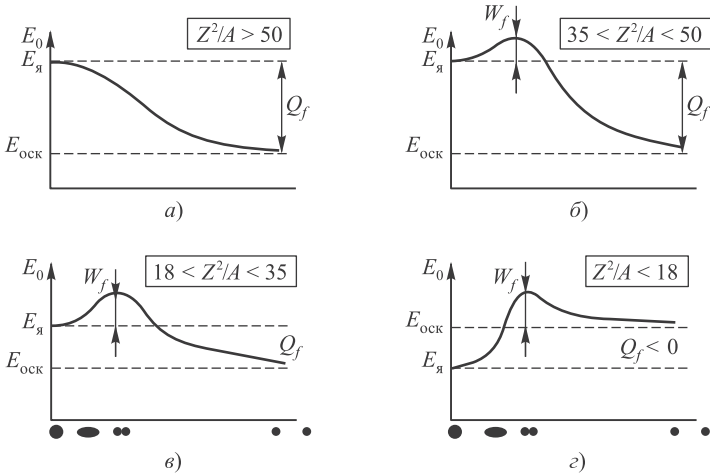
Если дефект массы  $\Delta_\epsilon M$  положителен, то при деформации масса ядра уменьшается, т. е. это энергетическое состояние более выгодно, что позволяет деформации развиваться дальше, пока ядро не развалится на меньшие осколки. Это происходит очень быстро, так что такие ядра просто не существуют в природе. Если же  $\Delta_\epsilon M < 0$ , то деформация увеличивает массу, и ядро стремится ее уменьшить, возвращаясь к сферической форме.

С помощью (44.62) критерий стабильности ядер  $\Delta_\epsilon M < 0$  можно записать в виде

$$\frac{Z^2}{A} < \frac{0,0072}{0,0001426} \approx 50. \quad (44.63)$$

Для элементов в начале и середине таблицы Менделеева это условие заведомо выполнено: например, для  ${}^{56}_{26}\text{Fe}$  получается значение  $Z^2/A = 12,1$ . Для стабильного ядра урана  ${}^{238}_{92}\text{U}$  имеем  $Z^2/A = 92^2/238 = 35,6$ . Для самого тяжелого из известных элементов оганесона ( ${}^{118}_{294}\text{Og}$ ) (см. сноску на с. 510) получаем  $Z^2/A = 118^2/294 = 47,4$ .

Для наиболее устойчивых ядер мы нашли связь между числом  $Z$  протонов и полным числом  $A$  нуклонов. Подставляя (44.17) в (44.63), находим предельные значения  $A_{\max} = 390$ ,  $Z_{\max} = 140$ . Конечно, такого элемента нет в природе и он никогда не будет создан искусственно. Подчеркнем еще раз: при нарушении условия (44.63) ядра просто не могут существовать, мгновенно разваливаясь на меньшие объекты. Характерное расстояние для ядерных сил — это радиус протона  $r_p$ , характерное время для таких типично ядерных процессов определяется временем, за которое сигнал успевает пройти «ядерное» расстояние:  $\tau_{я} \sim r_p/c \sim 10^{-23}$  с.



**Рис. 44.9.** Поведение энергии покоя (массы) ядра в зависимости от его деформации: а) нестабильные ядра (разваливающиеся за время порядка  $10^{-23}$  с); б) ядра, способные к спонтанному делению (потенциальный барьер  $W_f$ , отделяющий недеформированное состояние ядра от деформированного, невысок, и возможно туннелирование); в) ядра с высоким потенциальным барьером  $W_f$ , практически стабильные относительно спонтанного деления; г) стабильные ядра

Итак, мы убедились, что для всей таблицы Менделеева ядра устойчивы относительно малых деформаций. Однако представим себе, что ядро все-таки развалилось на два (для простоты равных) осколка. Посмотрим, как изменится при этом масса системы. Дефект массы в таком процессе равен:  $\Delta M(Z, A) = M(Z, A) - 2M(Z/2, A/2)$ , где  $M(Z, A)$  выражается формулой (44.20). Нетрудно получить, что

$$\begin{aligned} \Delta M(Z, A) &= 0,0180A^{2/3}(1 - 2^{1/3}) + 0,000713 \frac{Z^2}{A^{1/3}}(1 - 2^{-2/3}) = \\ &= -0,00468A^{2/3} + 0,000264 \frac{Z^2}{A^{1/3}}. \end{aligned} \quad (44.64)$$

Ядро стабильно относительно развала на два осколка при  $\Delta M(Z, A) < 0$ , откуда следует:

$$\frac{Z^2}{A} < \frac{0,00468}{0,000264} \approx 18. \quad (44.65)$$

Видно, что критерий стабильности (44.65) нарушается уже для элементов, следующих за ниобием  ${}_{41}^{93}\text{Nb}$  в таблице Менделеева.

Мы получили в какой-то степени противоречивый результат: с одной стороны, ядра с  $Z^2/A < 50$  устойчивы относительно деформаций, с другой — ядра с  $Z^2/A \geq 18$  все-таки могут разваливаться на осколки. Объяснение следует из рисунка 44.9, где схематически изображена зависимость энергии покоя (массы) ядра от степени его деформации. Стадии деформации ядра символически изображены под осью абсцисс: начальная сферическая и слегка деформированная эллипсоидальная формы, критическая деформация образования двух меньших

ядерных «капель», соединенных узкой перемычкой, и финальная стадия, когда осколки удалены на большое расстояние друг от друга.

Случай *a* соответствует ядрам, не удовлетворяющим критерию стабильности (44.63); такие ядра, как мы уже обсуждали, в высшей степени неустойчивы и разваливаются за время порядка  $10^{-23}$  с. Разность начальной и конечной энергий  $Q_f$  — это энерговыделение в процессе деления, имеющее своим источником дефект массы.

В трех случаях *b–г* кривая энергии возрастает при малых деформациях, поскольку выполнен критерий (44.63). Случай *г* соответствует ядрам, стабильным относительно процесса деления. В недеформированном состоянии они находятся в абсолютном минимуме по энергии: конечная масса осколков превышает массу начального ядра, т.е. энерговыделение отрицательно. Такие ядра удовлетворяют критерию стабильности (44.65), их распады не могут происходить самопроизвольно.

Два других случая (*b–в*) описывают ту самую противоречивую ситуацию, с которой мы разбираемся. При малых деформациях масса ядра увеличивается, т.е. состояние ядра устойчиво. При больших же деформациях ядро разваливается, его масса уменьшается и энерговыделение положительно. В недеформированном состоянии ядро находится в локальном минимуме энергии, но как скоро оно сможет выйти из него, чтобы попасть в глобальный минимум (т.е. развалиться на осколки) — вот в чем вопрос.

Выходу из локального минимума препятствует потенциальный барьер, высота которого  $W_f$  называется **барьером деления**. С ростом  $Z^2/A$  барьер деления быстро уменьшается: при  $Z^2/A \approx 32$  (золото, свинец)  $W_f \approx 40\text{--}50$  МэВ, но уже при  $Z^2/A \approx 35\text{--}36$  (для элементов вблизи урана и тория)  $W_f \approx 6\text{--}8$  МэВ. Энергия деления  $Q_f$  растет с ростом  $Z^2/A$ . Поскольку для элементов с  $Z^2/A < 35$  барьер велик, то такие ядра практически стабильны относительно деления. Этому случаю соответствует рис. 44.9, в.

Если барьер невелик, как показано на рис. 44.9, б, то даже устойчивое к малым деформациям ядро может «просочиться» через потенциальный барьер (квантовомеханическое туннелирование), т.е. развалиться на две меньшие части. Такое **спонтанное деление** ядер урана впервые наблюдали К. А. Петржак и Г. Н. Флеров (1940). Вероятность туннелирования сильно зависит от размера и высоты барьера. Так, время жизни для спонтанного деления урана очень велико ( $\approx 10^{16}$  лет), для других ядер оно быстро уменьшается с ростом  $Z^2/A$  вплоть до десятых долей секунды у «100-х» элементов.

Вывести ядро из локального минимума (сферическая форма) можно также с помощью какого-то внешнего воздействия (бомбардировка нейтронами), сообщаящего ядру дополнительную энергию. Критерием возможности такого процесса служит сравнение высоты барьера с энергией связи нейтрона  $\varepsilon_n$  в ядре. При  $W_f \leq \varepsilon_n$  нейтрон, не испытывающий кулоновского отталкивания, легко проникает в ядро  ${}^A_ZX$  и захватывается им. Возникает промежуточное ядро — изотоп  ${}^{A+1}_ZX$  первоначального элемента, выделяется энергия связи нейтрона, и этой энергии уже достаточно для деления промежуточного ядра. Если взять полученное выше выражение (44.21) и найти энергии связи ядер  $E_{\text{св}}(Z, A+1)$  и  $E_{\text{св}}(Z, A)$ , то их разность и даст нам энергию связи нейтрона  $\varepsilon_n$ . Для золота и урана получаются



**Таблица 44.5.** Энергия связи нейтрона  $\varepsilon_n$ , барьер деления  $W_f$  и возможность деления тепловыми нейтронами некоторых ядер

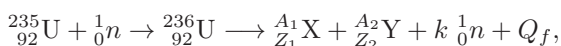
Ядро мишени	Промежуточное ядро	$\varepsilon_n$ (МэВ)	$W_f$ (МэВ)	Деление тепловыми нейтронами
$^{235}\text{U}$	$^{236}\text{U}$	6,5	5,2	да
$^{238}\text{U}$	$^{239}\text{U}$	4,8	5,7	нет
$^{239}\text{Pu}$	$^{240}\text{Pu}$	6,4	4,8	да
$^{243}\text{Am}$	$^{244}\text{Am}$	5,5	5,8	нет

примерно одинаковые значения  $\varepsilon_n \approx 6\text{--}7$  МэВ. Барьер деления для золота и соседних элементов почти на порядок выше, и потому они нейтронами не делятся. При захвате нейтрона такое ядро приходит в возбужденное состояние, но не делится, а «сбрасывает» лишнюю энергию, излучая  $\gamma$ -кванты или иные частицы (т.е. претерпевая  $\alpha$ - или  $\beta$ -распад). Для урана же и близких к нему элементов в таблице Менделеева барьер деления получился примерно таким же, как и энергия связи нейтрона. Поэтому элементы с  $Z^2/A \geq 35$  имеют  $W_f \leq \varepsilon_n$  и легко делятся быстрыми нейтронами с энергиями порядка 1 МэВ, а некоторые из них делятся даже тепловыми нейтронами с практически нулевой энергией. В таблице 44.5 приведены данные по энергии связи нейтрона и барьеру деления некоторых ядер. Из нее следует, например, что наиболее распространенный в природе изотоп урана  $^{238}\text{U}$  не делится тепловыми нейтронами. Однако быстрые нейтроны с энергией по крайней мере 1,3 МэВ приводят к делению и таких ядер.

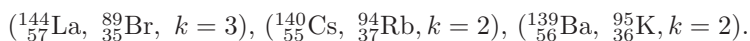
## Цепная реакция. Атомные бомбы и реакторы

В 1938 г. немецкие ученые О. Ган и Ф. Штрассман осуществили деление ядер урана при облучении их нейтронами. Природный уран содержит три изотопа:  $^{238}\text{U}$  (99,27%),  $^{235}\text{U}$  (0,72%),  $^{234}\text{U}$  (0,01%), т.е. приблизительно 1 атом  $^{235}\text{U}$  приходится на 140 атомов  $^{238}\text{U}$ .

Оказывается, что  $^{235}\text{U}$  при поглощении медленного (теплового) нейтрона с энергией  $\varepsilon_n \approx 0,025$  эВ может делиться по следующей схеме:



где осколками X, Y чаще всего бывают пары ядер:



Здесь указано также число  $k$  вылетающих нейтронов.

В своих опытах Ган и Штрассман, обстреливая уран нейтронами, обнаружили в продуктах реакции барий. Ввиду военного времени Нобелевская премия по химии за 1944 г. была присуждена Гану «за открытие расщепления тяжелых ядер» только в 1945 г., а получил он ее в 1946 г., так как был интернирован в Великобританию вместе с группой немецких специалистов по ядерной физике.

При делении промежуточного ядра  $^{236}\text{U}$  с наибольшей вероятностью образуются осколки, массы которых относятся как 2 : 3. Удельная энергия связи осколков примерно на 1 МэВ больше, чем у исходного ядра. Таким образом, процесс деления сопровождается выделением большого количества энергии. Но особо важным является тот факт, что при делении каждого ядра вылетает несколько нейтронов. В среднем на каждый акт деления  $^{235}\text{U}$  приходится  $k = 2,46$  выделившихся нейтронов, при делении плутония  $^{239}\text{Pu}$  это число больше:  $k = 2,91$ . Образующиеся при делении осколки (новые ядра) остаются радиоактивными и претерпевают цепочку бета-распадов, сопровождающихся испусканием гамма-излучения.

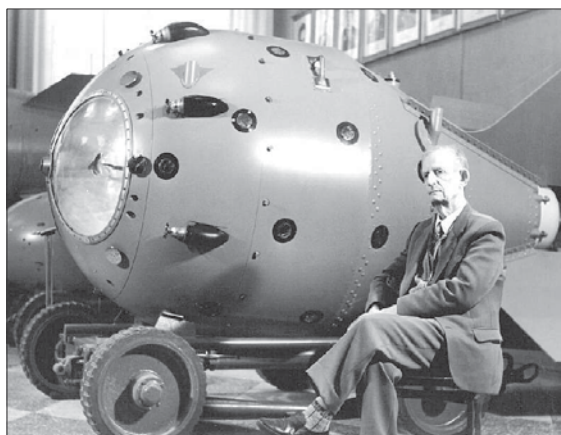
Испускание нескольких нейтронов при делении ядра делает возможной *цепную ядерную реакцию*. Она может осуществляться двумя способами. Первый способ заключается в выделении из природного урана *делящегося изотопа*  $^{235}\text{U}$  или получении промышленным путем  $^{239}\text{Pu}$ . Испускаемые при делении нейтроны любых энергий могут вызвать новые деления таких ядер урана или плутония с испусканием новых нейтронов, так что количество нейтронов, образующихся в каждом поколении, нарастает в геометрической прогрессии. Однако если масса делящегося изотопа меньше определенного критического значения, то большинство испущенных нейтронов вылетает наружу (за пределы так называемой активной зоны), не вызывая деления. В этом случае цепная реакция не возникает. Так происходит потому, что объем шара, изготовленного из делящегося материала, растет как куб его радиуса, а поверхность — как квадрат радиуса. Число рождающихся нейтронов пропорционально объему тела, а число вылетающих наружу — площади поверхности. Отношение площади поверхности к объему  $S/V \propto 1/R$ . Поэтому относительная доля вылетающих наружу через поверхность нейтронов уменьшается с ростом массы (радиуса) делящегося вещества. Приведем оценки для критических масс используемых на практике изотопов<sup>6</sup>: для урана-235 критическая масса  $M_{\text{кр}} = 50$  кг ( $R_{\text{кр}} = 9$  см), для плутония-239  $M_{\text{кр}} = 11$  кг, для урана-233  $M_{\text{кр}} = 16$  кг. При массе, превышающей критическую, нейтроны быстро размножаются, и реакция приобретает взрывной характер. На этом основано действие атомной бомбы.

В земной атмосфере всегда имеется некоторое количество нейтронов, рожденных космическими лучами. Поэтому, чтобы вызвать взрыв, достаточно соединить части ядерного заряда в один кусок с массой, большей критической.

Первые американские атомные бомбы, разработанные в Лос-Аламосе, назывались «Мальш» (*Little Boy*) и «Толстяк» (*Fat Man*). В них быстрое сведение вместе нескольких частей заряда ядерного вещества, каждая из которых имела массу меньше критической, осуществлялось с помощью обычного взрывчатого вещества.

Бомба «Мальш» представляла собой длинный цилиндр, закрытый с обоих концов. В одном конце цилиндра находилась «цель», представляющая некоторое количество ядерного взрывчатого вещества меньше критической массы, а в другом конце помещался «снаряд» из урана-235 также меньше критической массы. За этим «снарядом» помещался заряд обычного взрывчатого вещества. Устройство «Мальша» было относительно простым, но не слишком удачным с точки зрения коэффициента использования ядерного горючего.

<sup>6</sup> Данные взяты из книги: Физ. энциклопедия. М.: Советская энциклопедия, 1983. (См. также оценку, данную в т. 1, разд. 14.1 настоящего учебника.)

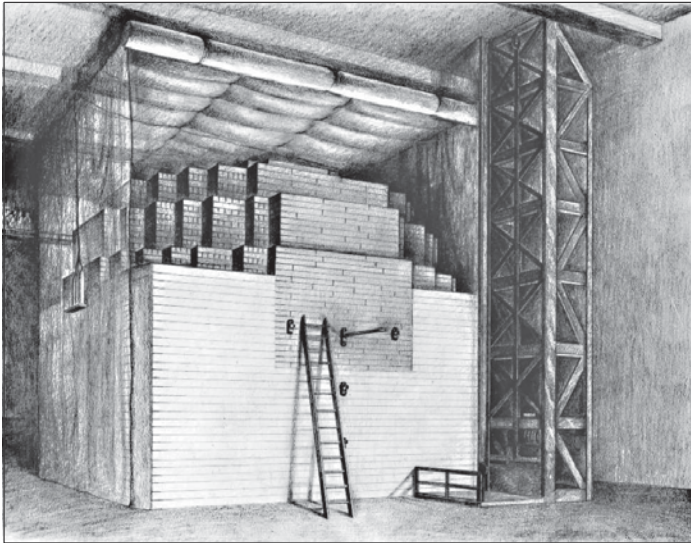


**Рис. 44.10.** Главный конструктор и научный руководитель КБ-11 трижды Герой Социалистического Труда академик Ю.Б. Харитон около макета РДС-1 в музее ядерного оружия (РФЯЦ—ВНИИЭФ)

Принцип, на котором основано действие «Толстяка», называется «имплозия», причем и принцип, и даже само это слово оставались засекреченными в течение многих лет после атомной бомбардировки Хиросимы и Нагасаки в августе 1945 г. В самых общих чертах примерная схема «Толстяка» выглядит следующим образом. В самом центре бомбы располагалась полая плутониевая сфера диаметром около 10 см. Внутри плутониевой сферы помещался бериллиевый шар, служивший источником нейтронов. Вокруг плутониевой сферы располагалась другая, очень большая сфера, состоящая из тридцати двух очень точно обработанных блоков взрывчатого вещества, выполненных в форме линз. В каждой из линз помещалось для большей надежности по два детонатора, соединенных в единую электрическую цепь. При одновременном срабатывании всех взрывателей (синхронизация осуществляется с большой точностью) мгновенная детонация химического взрывчатого вещества общей массой 2300 кг дает очень мощную взрывную волну, направленную в центр бомбы. Часть энергии этой волны фокусируется в месте расположения ядерного заряда. Все эти взрывные волны одновременно сойдутся в одной точке и сожмут полую плутониевую сферу с внутренним бериллиевым шаром до размеров теннисного мячика (диаметром около 6,5 см). Плотность плутониевого заряда достигнет критической величины, и под действием источника нейтронов начнется (и в течение одной микросекунды завершится) ядерная цепная реакция взрывного типа. Очень похоже была устроена первая советская атомная бомба РДС-1 (рис. 44.10), создававшаяся в КБ-11 в Арзамасе-16 (ныне г. Саров<sup>7</sup>) и испытанная в августе 1949 г. на полигоне в Семипалатинской области Казахской ССР.

Энергия, выделяющаяся при делении 1 г урана, эквивалентна энергии, выделяющейся при сгорании 2,5 т угля, и составляет 22 МВт · ч. Образующаяся при взрыве ударная волна несет избыточное давление  $p$ , и при  $p$  порядка нескольких десятых долей атмосферного давления происходят существенные поражения. Радиус поражения  $R$  приблизительно определяется из формулы  $p = E/R^3 = 10^5$  Па, где  $E$  — энергия,

<sup>7</sup>В 1967 г. КБ-11 было преобразовано во Всесоюзный научно-исследовательский институт экспериментальной физики (ВНИИЭФ), вошедший в структуру Министерства среднего машиностроения СССР. С февраля 1992 г. получил название Российский федеральный ядерный центр — Всероссийский научно-исследовательский институт экспериментальной физики (РФЯЦ—ВНИИЭФ). Является федеральным государственным унитарным предприятием ядерно-оружейного комплекса Государственной корпорации по атомной энергии «Росатом».



**Рис. 44.11.** «Чикагская поленица» (*Chicago pile-1*) — первый в мире ядерный реактор, построенный в США (1942) под руководством Э. Ферми; активная зона представляла собой кубическую решетку с шагом около 21 см, образованную слоями графитовых блоков, причем в каждом втором слое имелись отверстия для загрузки ядерного топлива (природного необогащенного урана); коэффициент размножения нейтронов был лишь чуть больше единицы ( $K = 1,0006$ )

выделяющаяся при ядерном взрыве. Для простоты использования этой формулы вспомним, что такое тротиловый эквивалент: при взрыве 1 кг тротила выделяется энергия 4,2 МДж (см. т. 1, ч. II, задача 10.42). Таким образом, взрыв одной килотонны тротила выделяет энергию  $4,2 \cdot 10^{12}$  Дж. Разделив обе части последнего равенства на эту величину, получим энергию взрыва  $E$ , выраженную в килотоннах тритилового эквивалента. Умножив дополнительно на  $10^9$ , выразим  $R$  в километрах:  $E, \text{ кт}/(R, \text{ км})^3 = 23,8$ , откуда

$$R = 0,35 \text{ км} \sqrt[3]{E/1 \text{ кт}}. \quad (44.66)$$

Например, при взрыве бомбы в 20 кт радиус поражения  $R \approx 1 \text{ км}$ .

**Задача 44.75.** Оценить, какое количество урана участвует в делении при взрыве 20-килотонной бомбы.

**Решение.** При делении урана выделяется энергия порядка 1 МэВ/нуклон, т. е.  $10^6 \times 1,6 \cdot 10^{-19} / 1,67 \cdot 10^{-27} \approx 10^{14}$  Дж/кг. Энергия же взрыва 20-килотонной бомбы составляет  $20 \times 4,2 \cdot 10^{12} = 0,84 \cdot 10^{14}$  Дж. Эти числа примерно совпадают, т. е. деление одного килограмма  $^{235}\text{U}$  выделяет такую же энергию, как и взрыв 20 килотонн тротила. ■

Усилия ученых были направлены не только на военные применения цепной реакции, но и на создание атомной (правильнее ядерной) энергетике. Первый в мире ядерный реактор был пущен в декабре 1942 г. в Чикагском университете под руководством итальянского физика Э. Ферми (рис. 44.11). В нашей стране реактор такого же типа

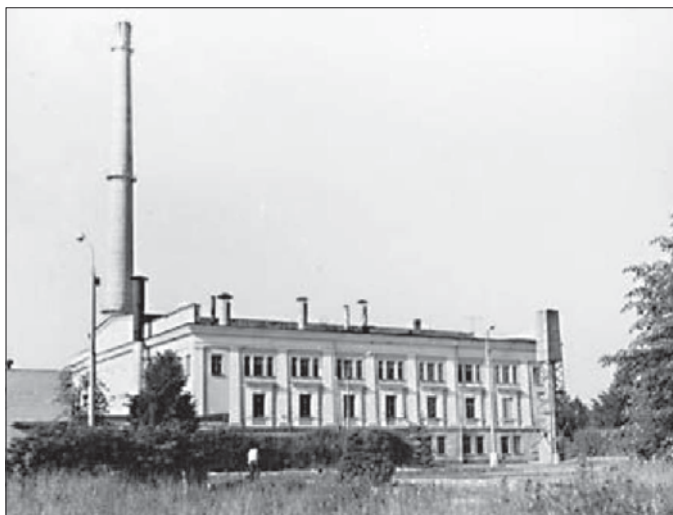


Рис. 44.12. Первая в мире АЭС (1954) в г. Обнинске Калужской области

был пущен под руководством И. В. Курчатова в декабре 1946 г. в Москве. Первая в мире атомная электростанция опытно-промышленного назначения (мощностью 5 МВт) была пущена в июне 1954 г. в одном из крупнейших научных центров СССР — Физико-энергетическом институте в Обнинске (рис. 44.12) Энергетический блок представлял собой водоохлаждаемый канальный уран-графитовый реактор, названный АМ-1. Начиная с 1956 г., Обнинская АЭС использовалась также как исследовательский полигон при создании новых, более мощных АЭС. Опыт ее эксплуатации полностью подтвердил технические и инженерные решения, предложенные специалистами отрасли, что позволило осуществить строительство и пуск в 1964 г. Белоярской АЭС электрической мощностью 300 МВт. «Ядерный котел» Обнинской АЭС работал безаварийно 48 лет, его заглушили навсегда в апреле 2002 г. Сейчас Обнинская АЭС превращена в отраслевой мемориальный комплекс, ее посещают организованные экскурсии и иностранные делегации.

В ядерных реакторах используется другой способ осуществления цепной реакции, нежели в атомной бомбе, а именно *управляемая цепная реакция*, при которой коэффициент размножения нейтронов  $K$  поддерживается на уровне чуть больше единицы. (Физический смысл величины  $K$  мы проанализируем позже.) В качестве активного вещества в реакторе может использоваться природный уран, обогащенный изотопом  $^{235}\text{U}$ . Нейтроны, испускаемые при делении ядер  $^{235}\text{U}$ , имеют в среднем энергию  $\varepsilon_n \sim 2$  МэВ, которая соответствует скорости  $v_n \sim 2 \cdot 10^7$  м/с. Эти *быстрые нейтроны* с большой вероятностью поглощаются ядрами неделящейся примеси, вследствие чего, не вызвав деления, выходят из игры и не порождают новых нейтронов. Чтобы предотвратить захват нейтронов ядрами неделящегося изотопа  $^{238}\text{U}$ , сравнительно небольшие блоки активного вещества размещают на некотором расстоянии друг от друга, а промежутки между блоками заполняют *замедлителем*, т. е. веществом, в котором быстрые нейтроны замедляются до тепловых скоростей (2200 м/с). Вероятность реакций

деления урана-235 тепловыми нейтронами почти в 200 раз больше вероятности захвата тепловых нейтронов ядрами урана-238. Замедление нейтронов осуществляется за счет упругого рассеяния. В этом случае энергия, теряемая замедляемой частицей, зависит от соотношения масс сталкивающихся частиц. Максимальное количество энергии теряется в случае, когда обе частицы имеют одинаковую массу (см. т. 1, раздел 4.7). С этой точки зрения идеальным замедлителем было бы вещество, содержащее обычный водород (например, вода — массы протона и нейтрона примерно одинаковы).

Однако вода оказалась непригодной в качестве замедлителя, потому что нейтроны поглощаются протонами, которые вступают с ними в реакцию  $p(n, \gamma)d$  образования дейтрона. Ядра замедлителя должны обладать малой вероятностью захвата нейтронов и большой вероятностью упругого рассеяния. Этому условию удовлетворяет дейтерий, а также ядра графита и бериллия. Для уменьшения энергии нейтронов, испускаемых при делении, до тепловых энергий  $\varepsilon_n \sim k_B T$  в тяжелой воде  $D_2O$  достаточно около 25 столкновений, в графите или бериллии — примерно ста столкновений.

Основные элементы каждого ядерного реактора независимо от его назначения и конструкции: ядерное топливо (а также приспособления для его обновления); замедлитель нейтронов; устройства (управляющие стержни) для регулирования хода цепной реакции (из материалов, сильно поглощающих нейтроны, например, бора, кадмия, гафния); предохранительные устройства, обеспечивающие защиту персонала от излучений; теплоноситель, отводящий избыток тепла; отражатель, уменьшающий потери нейтронов через поверхность активной зоны.

Управляемая цепная реакция является основой всех ядерных реакторов. В настоящее время существует много различных типов реакторов, которые используются как для исследовательских целей, так и для получения радиоизотопов, а также для производства энергии. По данным Международного агентства по атомной энергии (МАГАТЭ)<sup>8</sup> на 31.12.2015, в мире действовал 441 промышленный реактор, их общая электрическая мощность составила 382 855 МВт. В настоящее время используются следующие типы реакторов (указаны международная аббревиатура и общее число реакторов данного типа):

- водо-водяной реактор с обычной водой (под давлением) в качестве замедлителя и теплоносителя (*PWR* от англ. *Pressurized water reactor*) — 282;
- «кипящий» водо-водяной реактор, где пар генерируется непосредственно в активной зоне и направляется в турбину (*BWR* от англ. *Boiling Water Reactor*) — 78;
- тяжеловодный реактор, в котором замедлителем и теплоносителем является тяжелая вода  $D_2O$  (*PHWR* от англ. *Pressurised Heavy Water Reactor*) — 49;
- водо-графитовый реактор, использующий в качестве замедлителя графит, а в качестве теплоносителя обычную воду (*LWGR* от англ. *Light Water Graphite Reactor*) — 15 (все в России);

<sup>8</sup> *Nuclear Power Reactors in the World, 2016 Edition. IAEA, Vienna, 2016.*

- графито-газовые реакторы с газообразным теплоносителем (гелий, углекислый газ) и графитовым замедлителем (*GCR* от англ. *Gas-cooled reactor*) — 14 (все в Великобритании);
- реакторы-размножители (бридеры) (*FBR* от англ. *Fast breeder reactor*) — 3 (два в России на Белоярской АЭС, один в Китае).

Важнейшим параметром любого реактора является **коэффициент размножения нейтронов  $K$** , показывающий количество тепловых нейтронов, которые рождаются в каждом новом поколении в ходе цепной реакции. Его значение  $K_\infty$  в бесконечно большой среде можно определить по «формуле четырех сомножителей»:  $K_\infty = k\epsilon\phi\theta$ , где  $k$  — среднее число вторичных нейтронов (образующихся при поглощении нейтрона атомом ядерного топлива с последующим его делением),  $\epsilon$  — коэффициент размножения на быстрых нейтронах (т. е. увеличение числа нейтронов за счет деления  $^{238}\text{U}$ ),  $\phi$  — вероятность того, что нейтрон не будет захвачен ядром  $^{238}\text{U}$  и  $\theta$  — вероятность того, что тепловой нейтрон вызовет деление. Например, для графитовых реакторов на обогащенном (3–4%  $^{235}\text{U}$ ) естественном уране  $k = 2,46$ ,  $\epsilon = 1,03$ ,  $\phi\theta = 0,43$  и  $K_\infty = 1,09$ . Поскольку активная зона реального реактора имеет конечные размеры, коэффициент размножения равен  $K = K_\infty(1 - P)$ , где  $P$  — вероятность вылета нейтронов из активной зоны. Цепная реакция будет поддерживаться при условии  $K \geq 1$ , откуда для нашего примера находим  $P \leq 0,08$ . Это осуществляется лишь при определенных критических размерах активной зоны реактора — порядка 5–10 м. Это в 10–20 раз больше критического размера бомбы; соответственно масса активной зоны реактора будет больше массы бомбы в  $10^3$ – $10^4$  раз и составит десятки тонн. Первый ядерный реактор (см. рис. 44.11) имел коэффициент размножения нейтронов  $K = 1,0006$ .

Если все реакторы будут работать на обогащенном уране, т. е. на тепловых нейтронах, как указано выше, то при нынешнем уровне использования ядерной энергии существующие запасы урана будут израсходованы в течение нескольких десятилетий. По оценкам, коммерческие запасы  $^{235}\text{U}$  не превышают по своему энергетическому потенциалу запасы нефти и не могут кардинально решить энергетическую проблему, а добыча урана в России составляет 7% от общемировой.

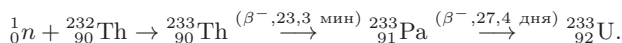
Проблему ядерного горючего можно преодолеть, если построить реакторы-размножители (*бридеры*, от англ. *to breed* — размножаться, плодиться). Такие реакторы дают больше топлива, чем потребляют. Как пример рассмотрим активную зону реактора, содержащую в качестве топлива плутоний-239 и природный уран-238. При делении ядра плутония-239 возникает в среднем 2,91 нейтронов. Один из этих быстрых нейтронов может вызвать деление другого ядра плутония-239, а второй может быть захвачен ядром природного урана-238 и привести к реакции:



(над стрелками указаны также периоды полураспада).

Таким образом, в результате захвата нейтрона ураном-238 образуются ядра плутония-239, способные к делению. Плутоний-239 радиоактивен, он претерпевает альфа-распад с образованием урана-235, однако период полураспада плутония весьма

велик:  $T_{1/2} = 24\,400$  лет. Другой вид ядерного топлива, которое можно получить искусственным путем в реакторе, — это уран-233, образующийся при облучении нейтронами тория-232:



Уран-233 претерпевает альфа-распад с периодом полураспада  $T_{1/2} = 162$  тыс. лет.

Основной параметр реактора-размножителя — это коэффициент воспроизводства, представляющий собой отношение числа атомов вторичного ядерного топлива к числу атомов первичного. Тот факт, что реакторы-размножители не использовались с начала развития ядерной энергетики, объясняется значительными техническими трудностями.

С одной стороны, в реакторах на быстрых нейтронах вследствие более интенсивного нейтронного потока предъявляются особо высокие требования к различным материалам и веществам (проблема радиационной стойкости материалов). Имеются трудности в создании материалов, используемых в тепловыделяющих элементах (ТВЭЛ). Использовать воду не представляется возможным, и в качестве теплоносителя применяют жидкий натрий.

С другой стороны, требуется разработка наиболее эффективных методов регенерации ядерного топлива, включающая выделение плутония-239. Причем срок наработки плутония в количестве, достаточном для удвоения реакторной мощности, желательно иметь не более 10–12 лет.

Самые мощные в мире бридеры впервые были построены в СССР: один реактор (БН-350, мощность 350 МВт) — на полуострове Мангышлак в г. Шевченко (1972, сейчас остановлен), другой — БН-600 с мощностью 600 МВт (1980), который до сих пор функционирует на Белоярской АЭС. По сообщениям прессы, в апреле 2016 г. бридер нового поколения БН-800 с проектной электрической мощностью 880 МВт вышел на 85% номинальной мощности, а в октябре 2016 г. был подписан приказ о введении энергоблока № 4 (БН-800) Белоярской АЭС в промышленную эксплуатацию. Проект БН-800, разработанный еще в 1983 г., с тех пор дважды пересматривался в свете новых требований к безопасности эксплуатации. Старейший американский журнал по энергетике «Power» присудил четвертому энергоблоку (БН-800) премию «Power Awards» за 2016 г. в номинации «Лучшие станции».

Бридеры строились также в Европе и США, но в настоящее время все они закрыты. Последний в Европе бридер — «Суперфеникс» во Франции мощностью 1,2 ГВт — производил электричество с 1986 по 1996 г.

Реакторы на быстрых нейтронах могут в пределах использовать весь природный уран, и в этом — преимущество бридеров. Энергетические запасы урана-238 в земной коре могли бы решить проблемы энергетики в следующем тысячелетии, но пока большие надежды на быстрое развитие бридеров не оправдываются.

Хотя на пути развития ядерной энергетики на основе реакций деления тяжелых элементов существует ряд препятствий, в том числе и психологического порядка (и многие опасения вполне резонны), ей пока не видно альтернативы. Уже сейчас 10,9% мирового производства электроэнергии приходится на долю атомных электростанций. Во многих странах АЭС поставляют более трети потребляемой энергии (во Франции — 76,3%). В таблице 44.6 приведены данные МАГАТЭ о производстве электроэнергии на АЭС в ведущих в этой отрасли странах.



**Таблица 44.6.** Мировые лидеры в производстве электрической энергии на ядерных реакторах, по данным Международного агентства по атомной энергии (МАГАТЭ) на 31.12.2015; в последней колонке указана доля АЭС в общем производстве электроэнергии в данной стране

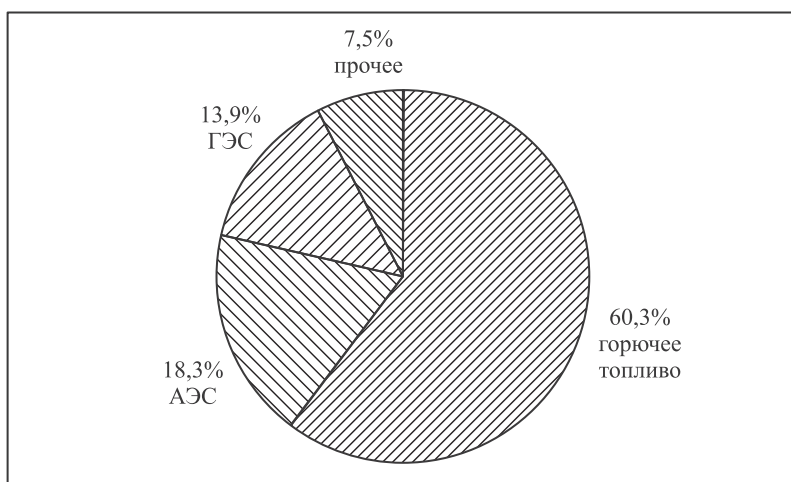
Страна	Число реакторов	Мощность (МВт)	Производство энергии (ТВт·ч)	%
США	99	99 185	796,0	19,5
Франция	58	63 130	419,0	76,3
Россия	35	25 443	182,8	18,6
Китай	31	26 774	161,2	3,0
Южная Корея	24	21 733	157,2	31,7
Канада	19	13 524	95,6	16,6
Германия	8	10 799	86,8	14,1
Великобритания	15	8 918	83,9	18,9
Украина	15	13 107	82,4	58,5
Испания	7	7 121	54,8	20,3
Швеция	10	9 648	54,5	34,3

На рисунке 44.13 приведены данные Международного энергетического агентства (*IEA*) о производстве электроэнергии в странах, входящих в Организацию экономического сотрудничества и развития (*OACD*), членами которой являются 35 стран — государства Евросоюза, США, Канада, Мексика, Япония, Австралия, Израиль, Новая Зеландия и др. (из постсоветских стран в *OACD* входят лишь Эстония с 2010 г. и Латвия с 2016 г.).

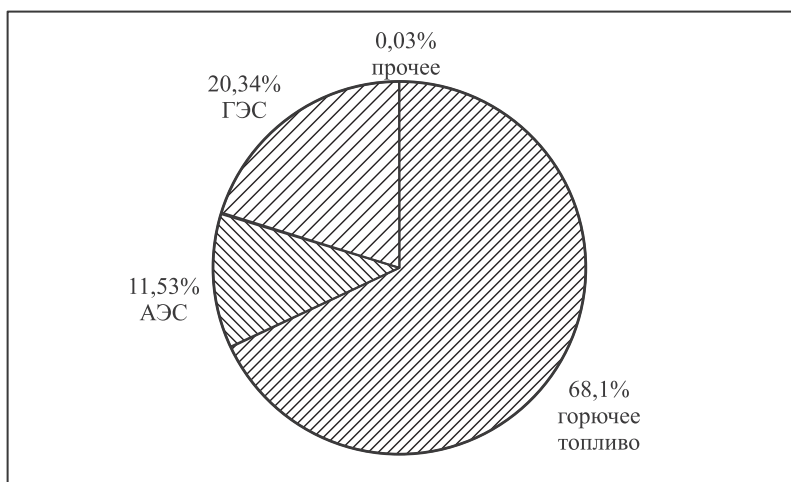
На рисунке 44.14 показаны аналогичные данные по России на 1.01.2016, приведенные Минэнерго на своем сайте (<http://minenergo.gov.ru/node/532>). В электроэнергетический комплекс России входят около 700 электростанций мощностью выше 5 МВт. Общая мощность электростанций единой энергосистемы России составила 235,30 ГВт. Видны различия в структуре производства энергии, связанные, очевидно, с географическими особенностями нашей страны. Как и во всем современном мире, львиная доля производства электроэнергии приходится на горючее топливо (в основном ископаемое).

**Задача 44.76.** Определить массовый расход  $dm/dt$  ядерного горючего  $^{235}\text{U}$  в реакторе АЭС. Тепловая мощность станции равна  $P = 10$  МВт. Принять, что в одном акте деления выделяется энергия  $Q = 200$  МэВ, а КПД станции равен  $\eta = 0,2$  (20%).

**Решение.** Если за время  $dt$  разделилось  $dN$  ядер урана, то выделенная энергия равна  $QdN$ , а мощность реактора с учетом его КПД составит  $P = \eta QdN/dt$ . Масса ядра урана равна примерно  $m_U = 235 \times 1,67 \cdot 10^{-27} = 3,92 \cdot 10^{-25}$  кг, т. е. искомый массовый



**Рис. 44.13.** Источники электроэнергии в странах *ОАСД*, по данным Международного энергетического агентства на 2015 г.; в позицию «прочие источники» отнесены геотермальные, солнечные, ветряные, приливные и прочие возобновляемые источники энергии



**Рис. 44.14.** Источники электроэнергии в России, по данным Минэнерго РФ на 1.01.2016

расход составляет  $dm/dt = m_U dN/dt = m_U P / (\eta Q)$ . В СИ  $Q = 200 \cdot 10^6 \times 1,60 \cdot 10^{-19} = 3,20 \cdot 10^{-11}$  Дж. Отсюда находим:

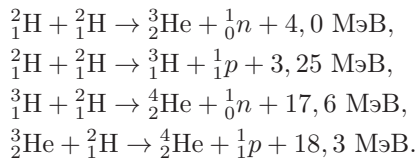
$$\frac{dm}{dt} = \frac{3,92 \cdot 10^{-25} \times 10^7}{0,2 \times 3,20 \cdot 10^{-11}} = 6,13 \cdot 10^{-7} \text{ кг/с} = 53 \text{ г/сут.} \blacksquare$$

## Термоядерная энергетика

Слияние легких ядер в одно ядро (ядерный синтез) сопровождается, как и деление тяжелых ядер, выделением огромных количеств энергии, причем для реакций синтеза происходит большее выделение энергии в расчете на один нуклон (или единицу массы топлива). При слиянии двух протонов и двух нейтронов в ядро гелия выделяется энергия 28,2 МэВ, а для 1 кг синтезированного гелия это составляет около 200 ГВт · ч, что в 10 раз превышает энерговыделение при делении урана.

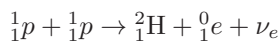
На первый взгляд методика осуществления этой реакции очень проста. Единственная сложность состоит в преодолении кулоновского отталкивания ядер при их сближении, поскольку протоны и другие легкие ядра положительно заряжены. Чтобы преодолеть потенциальный барьер, обусловленный кулоновским отталкиванием одноименно (положительно) заряженных частиц — протонов и других ядер, они должны обладать достаточной кинетической энергией, порядка 0,35 МэВ, что соответствует температуре  $T \approx 2$  млрд градусов. Однако вследствие туннельного эффекта, реакции синтеза могут протекать с заметной интенсивностью уже при температурах порядка нескольких сот миллионов кельвинов (для сравнения укажем, что температура на поверхности Солнца равна примерно 6000 К). Поскольку для синтеза ядер необходимы очень высокие температуры, этот процесс называется *термоядерной реакцией*.

Перечислим несколько основных реакций слияния ядер и приведем для них значения выделяющейся энергии:

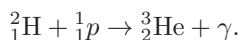


Выделяющейся в каждом акте синтеза энергии достаточно для того, чтобы участвующие в реакции ядра преодолевали кулоновское отталкивание. В последних двух реакциях — синтезе ядер дейтерия и трития — условия особенно благоприятны, так как взаимодействие ядер имеет резонансный характер. В этих реакциях энерговыделение в расчете на один нуклон в четыре раза превышает соответствующую величину для реакций деления тяжелых ядер.

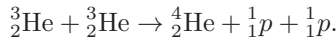
Синтез ядер водорода в ядра гелия, по-видимому, является источником энергии Солнца и звезд, температура в недрах которых  $\sim 10^7$ – $10^8$  градусов. Этот синтез может осуществляться двумя путями. При относительно низких температурах ( $\sim 10^7$  К) происходит протон-протонный цикл:



(напомним, что антиэлектрон  ${}^0_1e$  (другое обозначение  $e^+$ ) с зарядом +1 называется *позитроном*). Образовавшийся дейтрон сталкивается с протоном и объединяется в ядро гелия с испусканием  $\gamma$ -кванта:

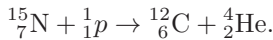
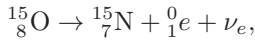
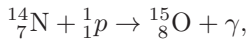
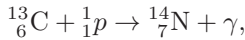
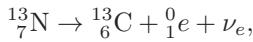
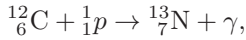


Последним звеном цикла является реакция



Итогом цикла является превращение четырех протонов в ядро гелия.

При более высоких температурах ( $T > 2 \cdot 10^7$  К) большей вероятностью обладает предложенный Г. Бете углеродный (или углерод-азотный) цикл:



Как и в протонном цикле, итогом углеродного цикла является исчезновение четырех протонов и образование одного ядра гелия, количество ядер углерода остается неизменным; эти ядра участвуют в реакции в роли катализатора.

Таким образом, топливом, дающим солнечную энергию, является обычный водород, а продуктом горения («золотой») — гелий. Оценки показывают, что водорода при неизменном выделении энергии хватит приблизительно на  $10^{11}$  лет, а возраст Солнца оценивается примерно в  $5 \cdot 10^9$  лет. В звездах, не сходных с нашим Солнцем, ядерные реакции, по-видимому, могут протекать иначе.

В земных условиях *неуправляемые термоядерные реакции* осуществляются при взрывах термоядерных (водородных) бомб. В качестве заряда в них используется смесь дейтерия и трития, в которой происходит неуправляемая самоподдерживающаяся термоядерная реакция синтеза взрывного типа. Запалом в водородной бомбе служит атомная бомба, при взрыве которой возникает температура порядка  $10^7$ – $10^8$  градусов. В течение микросекунд, за которые длится атомный взрыв, температура оказывается достаточно высокой для осуществления термоядерной реакции, которая высвобождает энергию и далее уже сама поддерживает температуру, необходимую для того, чтобы все термоядерное горючее успело «выгореть». Такие реакции имеют неконтролируемый характер.

Для осуществления *управляемых термоядерных реакций* необходимо создать и поддерживать в некотором объеме температуру выше сотни миллионов градусов, при которой вещество находится в состоянии *горячей плазмы* (голые ядра). Анализ баланса энергии в термоядерной реакции показывает, что для превышения энергетических затрат на создание и поддержание реагирующей плазмы должен выполняться так называемый *критерий Лоусона*: произведение плотности  $n$  плазмы (т. е. числа ионов в  $1 \text{ см}^3$ ) на время  $\tau$  реакции (иначе говоря, время удержания плазмы в реагирующем состоянии) должно быть больше некоторой величины, зависящей от температуры. Эта величина минимальна при температуре порядка  $10^8$  градусов, и критерий Лоусона выглядит как

$$n\tau \geq 2 \cdot 10^{20} \text{ с/м}^3. \quad (44.67)$$

Физический смысл критерия Лоусона состоит в том, что в термоядерном реакторе должно выделяться больше энергии, чем ее было затрачено для нагревания и удержания плазмы.

**Задача 44.77.** Пусть термоядерный реактор позволяет удерживать плазму в течение  $\tau = 980$  мс. Какой должна быть плотность плазмы  $n$  для осуществления термоядерной реакции? Сравнить  $n$  с плотностью  $n_g$  газа при нормальных условиях.

**Решение.** По критерию Лоусона находим:  $n \geq 2 \cdot 10^{20} / 980 \cdot 10^{-3} = 2,0 \cdot 10^{20} \text{ м}^{-3}$ . Рассмотрим один моль газа при нормальных условиях. Там содержится  $N_A = 6,01 \cdot 10^{23}$  частиц, занимающих стандартный молярный объем  $V_m = 22,46 \text{ л} = 22,46 \cdot 10^{-3} \text{ м}^3$ . Таким образом, плотность газа равна  $n_g = N_A / V_m = 2,7 \cdot 10^{25} \text{ м}^{-3}$ . Это больше плотности плазмы в 135 000 раз. ■

Итак, проблема управляемого термоядерного синтеза сводится, в сущности, к следующему: надо нагреть смесь дейтерия и трития до 100 миллионов кельвинов и удерживать в таком состоянии достаточно долго, пока выделяющаяся при термоядерном синтезе энергия не превысит все затраты. Достичь нужных значений параметра Лоусона  $n\tau$  можно двумя способами: увеличением времени удержания  $\tau$  (это можно сделать за счет термоизоляции плазмы магнитным полем — так называемое магнитное удержание) или за счет увеличения плотности плазмы  $n$  до таких значений, когда реакции синтеза успеют произойти раньше, чем плазменное облачко разлетится (инерционное удержание). При этом, очевидно, нагрев вещества, называемого мишенью, должен быть осуществлен за время, меньшее времени разлета, что требует от источников энергии большой мощности. При таком импульсном нагреве произойдет термоядерный микровзрыв.

Таким образом, критерий Лоусона можно выполнить, если нагреть очень плотную плазму ( $n \sim 10^{24}$  ионов/см<sup>3</sup>) до температур в сотни миллионов кельвинов, когда даже маленькое облачко с размерами порядка 1 мм будет разлетаться достаточно долго ( $\tau \sim 1$  нс). При таком подходе основной проблемой является сжатие вещества и его быстрый нагрев до требуемых температур, что можно надеяться осуществить с помощью лазеров и пучков релятивистских частиц. Термоядерные системы, основанные на этом принципе, называют системами с инерционным удержанием (т.е. со свободным разлетом плазмы). Предельным и уже реализованным случаем такой системы является взрыв водородной бомбы.

**Задача 44.78.** Таблетка термоядерного горючего сделана из смеси дейтерия и трития в равных долях. Плотность таблетки  $\rho_0 = 200 \text{ кг/м}^3$ . Лазерный импульс сжимает таблетку, увеличивая ее плотность в  $k = 1000$  раз. Какова плотность  $n$  получающейся плазмы и в течение какого времени ее надо удерживать, чтобы осуществить термоядерную реакцию?

**Решение.** Массы ядер дейтерия и трития примерно равны  $m_d = 2m_p = 3,34 \cdot 10^{-27} \text{ кг}$  и  $m_t = 3m_p = 5,01 \cdot 10^{-27} \text{ кг}$ . Поскольку дейтерий и тритий смешаны в равных пропорциях, средняя масса ядра в таблетке равна  $m = (m_d + m_t) / 2 = 4,18 \cdot 10^{-27} \text{ кг}$ . Плотность плазмы после сжатия находим теперь как  $n = 1000\rho_0 / m = 4,78 \cdot 10^{31} \text{ м}^{-3}$ . Согласно критерию Лоусона, плазму надо удерживать в течение времени  $\tau = 2 \cdot 10^{20} / 4,78 \cdot 10^{31} = 4,2 \cdot 10^{-12} \text{ с}$ , в продолжение которого таблетка должна оставаться сжатой, пока не начнется термоядерная реакция. ■

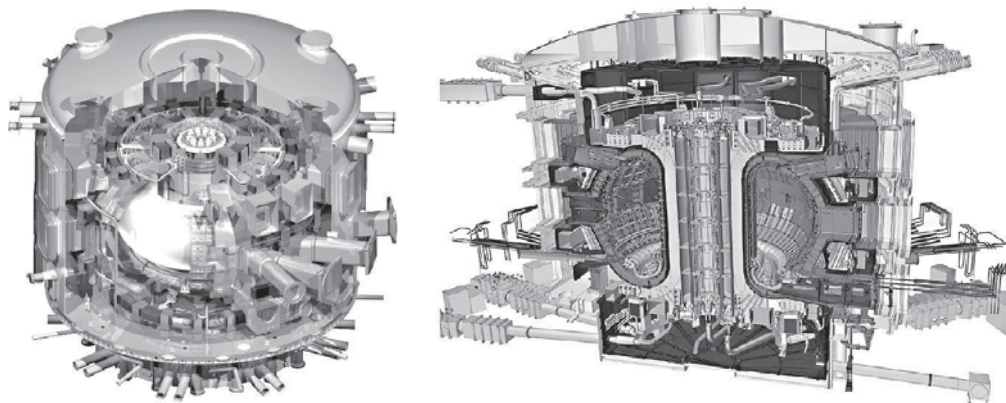


Рис. 44.15. Модель строящегося термоядерного реактора *ITER*

Задача исследований по инерционному удержанию сводится к уменьшению энергии взрыва до приемлемых значений и осуществлению серии управляемых микровзрывов. Чтобы удовлетворить критерию Лоусона при более низких плотностях плазмы и получить спокойное, почти стационарное «горение», необходимо найти способы эффективной термоизоляции плазмы. В принципе это можно сделать с помощью магнитных полей, предотвращающих контакт горячей плазмы со стенками термоядерного реактора. В системах с магнитным удержанием задача заключается в осуществлении длительной термоизоляции плазмы (например, до 1 с) при значениях плотности около  $10^{14}$  ионов/см<sup>3</sup>. В зависимости от структуры магнитных полей возможны различные виды ловушек для плазмы: *токамаки*, *стеллараторы*, *амбиполярные ловушки* и т. д. Схема токамака (аббревиатура от *Т*ороидальная *К*Амера с *М*Агнитной *К*атушкой) была предложена И. Е. Таммом и А. Д. Сахаровым (1951), этот проект получения управляемого термоядерного синтеза развивался в Институте атомной энергии им. И. В. Курчатова под руководством Л. А. Арцимовича, позже Б. Б. Кадомцева.

В токамаках плазма удерживается тороидальным внешним магнитным полем и полем тока, протекающего по плазменному шнуру. Поначалу эти установки существовали лишь в СССР, теперь исследовательские токамаки работают в различных странах мира. В наши дни в южной Франции 35 стран-участниц — Евросоюз, Индия, Китай, Россия, США, Южная Корея и Япония — осуществляют амбициозный проект, рассчитанный на несколько десятилетий: сооружение и эксперименты на прототипе промышленного термоядерного реактора ИТЭР (*ITER* от англ. *International Thermonuclear Experimental Reactor*, но также от лат. *iter* — путь). Европейский союз финансирует 45,6% бюджета проекта, остальные шесть стран вносят 9,1% каждая. Токamak *ITER* — гигантское сооружение высотой 60 м и массой 23 тыс. тонн (рис. 44.15). Планы сооружения *ITER* неоднократно пересматривались и уточнялись, на 2016 г. предполагалось, что строительство

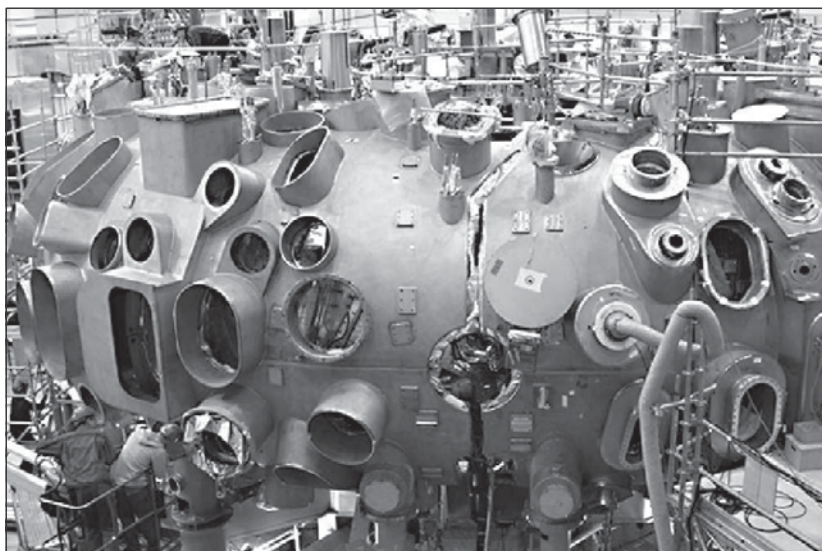


Рис. 44.16. Стелларатор «Вендельштайн»

завершится к 2021 г., а первая плазма будет получена в декабре 2025 г. Фотографии и обновляемая информация размещаются на сайте <https://www.iter.org/>.

Стелларатор (от лат. *stella* — звезда, название отражает цель достичь «звездных» температур плазмы) изобрел американский ученый Л. Спитцер (1950). Стелларатор отличается от токамака более сложной конструкцией, поскольку магнитная ловушка создается в нем лишь внешними катушками. Такая схема имеет свои плюсы и минусы; споры о том, что лучше, не прекращаются. В декабре 2015 г. в Германии на новом стеллараторе «Вендельштайн 7-X» (по названию горного массива в Баварских Альпах) получена первая тестовая плазма, в феврале 2016 г. начались эксперименты с водородной плазмой. «Вендельштайн» — самый мощный на 2016 г. стелларатор в мире (рис. 44.16). Его строительство велось 15 лет, затраты составили порядка 1 млрд евро. Реактор состоит из 50 сверхпроводящих ниобий-титановых катушек около 3,5 м в высоту и общим весом около 425 т, способных создавать магнитное поле индукцией 3 Тл, удерживающее плазму с температурой 60–130 млн градусов (что в несколько раз выше, чем температура в центре солнечного ядра). Объем плазмы может достигать  $30 \text{ м}^3$ , а ее плотность  $n = 3 \cdot 10^{20} \text{ м}^{-3}$ . При таких параметрах из критерия Лоусона (44.67) следует, что необходимое время удержания плазмы должно быть не меньше одной секунды. Одним из преимуществ стеллараторов является возможность работы в непрерывном режиме, в новой установке время запуска может составить до 30 мин непрерывной работы. Если физикам удастся достичь поставленной цели — удержать нагретую плазму в течение получаса, то, как они считают, откроется путь к сооружению промышленного стелларатора для коммерчески выгодного производства электроэнергии.

Даже если оптимистические прогнозы физиков оправдаются и управляемый термоядерный синтез будет осуществлен в следующем десятилетии, следует помнить о консервативности энергетической системы, которую невозможно перевести на новые виды топлива быстрее чем за время порядка 50 лет. Так что широкое практическое использование термоядерной энергетики — это, скорее всего, дело конца XXI века.

## Контрольные вопросы

1. Из каких составляющих построено атомное ядро? Перечислите их свойства.
2. Что такое изотопы? Приведите примеры.
3. Как вы думаете, чего в вашем теле больше — протонов или нейтронов?
4. Почему массовые числа для многих элементов в таблице Менделеева не целочисленны?
5. Откуда мы знаем о существовании сильного (ядерного) взаимодействия? Перечислите свойства ядерных сил.
6. Каков механизм возникновения ядерных сил между нуклонами?
7. Что такое дефект массы и энергия связи ядра ?
8. Нарисуйте примерный график зависимости удельной энергии связи ядра от атомного номера элемента и объясните принципиальную возможность получения полезного выхода энергии при превращении элементов.
9. Изложите основные идеи капельной модели ядра. Какие свойства ядер она объясняет?
10. Какие свойства ядер объясняет оболочечная модель ядра?
11. Сформулируйте закон радиоактивного распада. Что такое постоянная распада и период полураспада?
12. Укажите как можно больше различий между альфа-, бета- и гамма излучениями.
13. Почему скорость  $\alpha$ -распада сильно меняется от ядра к ядру, в то время как энергии вылетающих альфа-частиц близки друг к другу?
14. Почему энергия альфа-частиц при распаде данного ядра всегда постоянна, а энергия электронов при бета-распаде может принимать разные значения?
15. Объясните идею радиоуглеродного метода датировки ископаемых образцов органического происхождения. Перечислите преимущества и недостатки метода.
16. Укажите способ, как осуществить мечту средневековых алхимиков о превращении ртути в золото.
17. Объясните разницу между единицами рад, бэр и рентген.
18. Чем опасно  $\alpha$ -излучение, если от него так легко защититься?



19. Почему медленные нейтроны, имеющие небольшую кинетическую энергию, при взаимодействии с ядрами некоторых элементов вызывают их деление?
20. Почему практическое применение нашло деление ядер именно нейтронами, а не протонами?
21. Почему при замедлении быстрых нейтронов используют вещества, содержащие легкие элементы?
22. Обсудите сравнительные преимущества и недостатки (в том числе загрязнение окружающей среды и безопасность) различных способов получения энергии: сжигание ископаемого горючего, деление ядра, ядерный синтез.

## Глава 45

# Основные представления физики элементарных частиц

Размышлениям об основных «кирпичиках» природы, из которых построено все видимое разнообразие окружающего нас мира, ученые предавались со времен Античности. В разное время на роль этих «кирпичиков» претендовали различные физические объекты. Хорошим кандидатом со времен Демокрита был атом. Однако шло время, обнаруживались все новые виды атомов, и становилось ясно, что атом — это не самый фундаментальный объект природы. Получив в распоряжение квинтэссенцию экспериментальных данных по химии того времени — периодическую таблицу Д. И. Менделеева, ученые пытались найти более общий принцип, по которому она построена, нежели атомный вес и сходство химических свойств элементов. Как мы уже знаем, для объяснения химических свойств атомов и их расположения в таблице Менделеева было достаточно трех элементарных частиц — протона, нейтрона и электрона — и одного фундаментального взаимодействия протона с электроном — электромагнитного. Ядро построено из протонов и нейтронов, электрон связан с ядром электромагнитными силами, вся эта система описывается квантовой механикой — вот основные достижения физики 1920–1930 гг. Но природа оказалась устроенной не так просто. В этой главе мы кратко изложим современные представления о фундаментальных составляющих окружающего нас вещества и о взаимодействиях между ними.

### 45.1 Фундаментальные взаимодействия

В настоящее время физикам известно четыре типа фундаментальных взаимодействий: *электромагнитное*, *слабое*, *сильное*<sup>1</sup> и *гравитационное*. Гравитационное взаимодействие универсально, в нем участвуют все без исключения частицы. Электромагнитному взаимодействию подвержены частицы, имеющие

---

<sup>1</sup>Сейчас термин *сильное взаимодействие* чаще всего применяется к взаимодействию цветовых зарядов кварков (см. разд. 45.2), но ранее так называли взаимодействие нуклонов и мезонов, которое мы назовем *сильным ядерным взаимодействием*.

электрический заряд. Частицы, участвующие в сильных ядерных взаимодействиях, называются **адронами** (от греч. *hadros* — крупный, большой). Адроны — самый многочисленный отряд элементарных частиц: вместе с протоном, нейтроном и пи-мезонами к нему относятся еще несколько сот частиц, обнаруженных в экспериментах на ускорителях. Элементарные частицы, не участвующие в сильных ядерных взаимодействиях, называются **лептонами** (от греч. *leptos* — легкий). Лептонов всего шесть: электрон  $e^-$  со своими «двоюродными братьями» мюоном  $\mu^-$  и тау-лептоном  $\tau^-$  и соответствующие им нейтрино — электронное  $\nu_e$ , мюонное  $\nu_\mu$  и тау-нейтрино  $\nu_\tau$ . Плюс, конечно, их античастицы  $e^+$  (позитрон),  $\mu^+$ ,  $\tau^+$ ,  $\bar{\nu}_e$ ,  $\bar{\nu}_\mu$  и  $\bar{\nu}_\tau$  (с некоторыми из этих античастиц мы уже встречались в разд. 44.5).

Электромагнитное и гравитационное взаимодействия являются **дальнодействующими**, порождаемые ими силы между объектами простираются на космические расстояния, убывая по закону обратных квадратов  $F \propto 1/r^2$ . Слабое и сильное ядерное взаимодействия заметны на очень малых расстояниях, они убывают очень быстро (экспоненциально). Сильное (цветовое) взаимодействие, наоборот, растет с увеличением расстояния, за счет чего кварки прочно заперты внутри адронов. Три фундаментальных взаимодействия описываются соответствующими **квантовыми теориями поля**, в которых силы между объектами возникают вследствие обмена квантами данного поля. Квантовая теория гравитации пока не построена.

## Квантовая электродинамика

Первой квантовой теорией поля стала **квантовая электродинамика (КЭД)**. Сначала появилось обобщение квантовой механики: нерелятивистское уравнение Шрёдингера была заменено на **релятивистское уравнение Дирака** (1928), из которого автоматически следовало, в частности, существование у электрона спина  $1/2$  и магнитного момента. Уравнение Дирака предсказало также существование **позитрона** — античастицы электрона с положительным элементарным зарядом. В 1933 г. Э. Шрёдингер и П. Дирак были удостоены Нобелевской премии «за открытие новых продуктивных форм атомной теории». В 1932 г. позитрон был открыт в космических лучах американским физиком К. Д. Андерсоном (Нобелевская премия в 1936 г.).

Из уравнений Максвелла для электромагнитного поля следовало, что оно может быть представлено как набор связанных осцилляторов, распределенных по всему пространству. Переход к нормальным координатам позволил рассмотреть такую систему квантовомеханически как набор осцилляторов с непрерывно распределенными частотами  $\omega$ . Каждый такой осциллятор имеет уровни энергии  $E_\omega = \hbar\omega(n + 1/2)$ . Возбужденное  $n$ -е состояние с энергией  $\hbar\omega n$  интерпретируется как наличие  $n$  фотонов, соответствующих данной моде колебаний с частотой  $\omega$ . У фотонов, как следует из квантовой теории электромагнитного поля, спин равен единице. Однако возможны лишь два спиновых состояния, когда спин фотона направлен вдоль его импульса и против него. Для частиц, движущихся со скоростью света, состояние с нулевой проекцией спина на направление импульса

невозможно. Это квантовое свойство соответствует поперечности электромагнитных волн в классической электродинамике.

Уравнение Дирака допускало введение слагаемого, описывающего взаимодействие квантованного электромагнитного поля с релятивистскими квантовыми частицами — электронами и позитронами. Эта теория и называется квантовой электродинамикой (КЭД). Она с успехом описывает электромагнитные взаимодействия и других заряженных частиц (помимо электрона и позитрона). Основы КЭД — квантование электромагнитного поля и его взаимодействие с заряженными частицами — были заложены еще в конце 1920-х гг. Дираком, но при вычислениях возникали бесконечные выражения, бороться с которыми научились лишь на рубеже 1940-х — 1950-х гг., когда эта теория приняла современную форму. В 1965 г. японец С. Томонага и американцы Дж. Швингер и Р. Фейнман получили Нобелевскую премию «за фундаментальные работы по квантовой электродинамике, имевшие глубокие последствия для физики элементарных частиц» («за бортом», к сожалению, остался четвертый основоположник КЭД англо-американский физик Ф. Дайсон). Создание квантовой электродинамики завершило период развития собственно электромагнетизма — от классической к квантовой теории. Согласно КЭД, основным актом является испускание заряженной частицей фотона — кванта электромагнитного поля. Испущенный фотон виртуален — он быстро будет поглощен другой заряженной частицей, т. е. произойдет второй акт взаимодействия. В результате этих двух основных актов заряженные частицы будут притягиваться (отталкиваться) друг от друга. В каждом акте сохраняется электрический заряд, так что сами фотоны электрически нейтральны и не взаимодействуют друг с другом. Это проявляется в линейности уравнений электродинамики и справедливости принципа суперпозиции.

Наглядный метод описания взаимодействий дают *диаграммы Фейнмана*. Электроны и позитроны изображаются на них сплошными линиями, которые вследствие закона сохранения электрического заряда должны быть непрерывными. Фотоны изображаются волнистыми линиями. Вершины, где встречаются две электронные и одна фотонная линии, соответствуют элементарному акту взаимодействия. Линии между двумя вершинами соответствуют виртуальным (ненаблюдаемым) частицам; внешние линии диаграмм описывают реальные частицы. Стрелу времени на рисунке 45.1 и далее следует считать направленной слева направо.

Диаграммы Фейнмана — это не просто картинки: каждому элементу (внешним и внутренним линиям частиц и вершинам) соответствует определенный множитель, а диаграмме в целом — сложный многократный интеграл. В каждой вершине соблюдается закон сохранения энергии и импульса, но виртуальные частицы отличаются от реальных тем, что для них  $E^2 - p^2c^2 \neq m^2c^4$ , т. е. не выполняется обычное соотношение между энергией и импульсом (как говорят, виртуальные частицы находятся *вне массовой поверхности*). Сумма вкладов всех диаграмм дает амплитуду того или иного процесса, а квадрат амплитуды определяет его вероятность (сечение).

Расчеты в КЭД проводят по теории возмущений, т. е. последовательно учитывая диаграммы со все большим числом вершин. Каждая вершина вносит

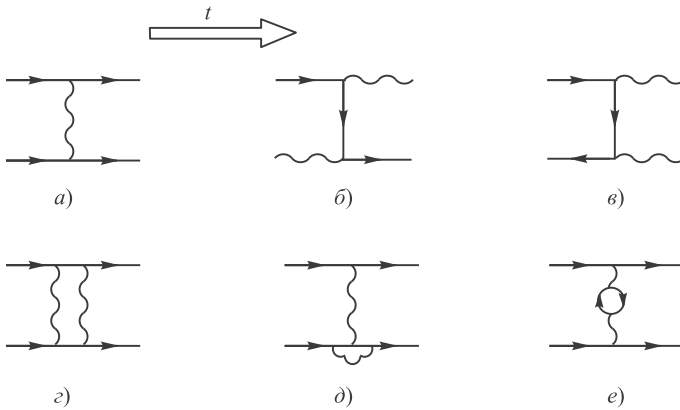


Рис. 45.1. Примеры диаграмм Фейнмана

множитель  $e$ , так что паре вершин соответствует квадрат элементарного электрического заряда, а точнее — безразмерная постоянная тонкой структуры  $\alpha_{em} = e^2/(4\pi\epsilon_0\hbar c) \approx 1/137$ . Малость этого параметра позволяет пользоваться процедурой разложения по его степеням.

Верхний ряд диаграмм на рис. 45.1, *a–в* описывает ряд процессов в низшем порядке теории возмущений: *a*) рассеяние электрона на электроне; *б*) рассеяние фотона на электроне (эффект Комптона); *в*) аннигиляцию электрона и позитрона в два фотона. На этой диаграмме (рис. 45.1, *в*) позитрону соответствует электронная линия, направленная против стрелы времени (на таком языке позитрон — это электрон, движущийся из будущего в прошлое).

Нижний ряд диаграмм на рис. 45.1, *г–е* представляет собой пример поправок следующего порядка к рассеянию электрона на электроне: *г*) обмен двумя фотонами; *д*) так называемая радиационная поправка к вершине; *е*) поправка к фотонной линии, описывающая распад виртуального фотона на виртуальную электрон-позитронную пару и ее последующую аннигиляцию в виртуальный фотон.

Заметим, что реальные электрон и позитрон не могут аннигилировать в один фотон. Действительно, этот процесс можно рассмотреть в системе центра масс электрона и позитрона. В ней полный импульс равен нулю и, стало быть, по закону сохранения импульса должен быть равен нулю полный импульс родившихся фотонов. Для двух фотонов это возможно (они просто имеют одинаковую энергию и разлетаются в противоположные стороны). Но при аннигиляции в один реальный фотон тот должен родиться с нулевым импульсом, т.е. в состоянии покоя, а это невозможно для частиц нулевой массы.

В 1954 г. американские физики Янг Чжэньнин и Р. Миллс предложили теорию, которая около десяти лет рассматривалась как изощренное математическое упражнение. Поясним их идею на самом простом примере. Как известно, волновую функцию  $\psi(\vec{r})$  квантовой механики можно умножить на фазовый множитель  $\exp(i\alpha)$ , не меняющий ее нормировку. Ни сами уравнения квантовой механики, ни следующие из них результаты при этом не меняются — говорят, что уравнения Шрёдингера и Дирака инвариантны относительно **глобального калибровочного преобразования**  $\psi(\vec{r}) \rightarrow \psi(\vec{r})e^{i\alpha}$ . Физическим следствием этой симметрии системы является закон сохранения электрического

заряда. Янг и Миллс заметили, что подобное преобразование волновой функции можно выполнить локально, т.е. считать фазу  $\alpha$  функцией радиуса-вектора. Уравнение Дирака для свободного электрона не будет инвариантным относительно **локального калибровочного преобразования**  $\psi(\vec{r}) \rightarrow \psi(\vec{r})e^{i\alpha(\vec{r})}$ , но станет таковым, если в него включить взаимодействие электрона с так называемым **полем Янга—Миллса**. Сам вид этого взаимодействия и даже уравнения для свободных полей Янга—Миллса почти однозначно следуют из свойств симметрии системы. В обсуждаемом простейшем случае симметрии относительно умножения волновой функции на фазовый множитель уравнения для полей Янга—Миллса совпадают с уравнениями Максвелла для электромагнитного поля и взаимодействия с ним движущегося электрического заряда. Иначе говоря, уравнения КЭД почти автоматически следуют из требования локальной калибровочной инвариантности релятивистского уравнения Дирака для свободного электрона.

Очевидно, что два последовательных калибровочных преобразования КЭД коммутируют (не зависят от порядка выполнения) — такая симметрия называется **абелевой**). Как следствие кванты абелевых полей Янга—Миллса (фотоны) не взаимодействуют друг с другом, уравнения таких полей линейны, и для них справедлив принцип суперпозиции.

Как выяснилось в дальнейшем, квантовые поля всех других фундаментальных взаимодействий также являются полями Янга—Миллса, соответствующими тем или иным группам симметрии рассматриваемых систем. Кванты полей Янга—Миллса — бозоны с нулевой массой и спином  $s = 1$  (здесь и далее спин указывается в единицах постоянной Планка  $\hbar$ ).

## Сильные ядерные взаимодействия

При проникновении в глубь атома, при попытках понять свойства атомных ядер, ученые столкнулись с ранее не известными силами, связывающими протоны и нейтроны в ядрах. Это **сильное ядерное взаимодействие** между нуклонами переносится  $\pi$ -мезонами. Мы уже говорили, что радиус  $R$  любого взаимодействия определяется массой  $m$  его переносчика:  $R \sim \hbar/(mc)$ . Масса фотона равна нулю, и потому электромагнитные силы являются дальнедействующими. Сильные ядерные взаимодействия имеют радиус  $R \sim 10^{-15}$  м, что соответствует массе промежуточных частиц около  $140 \text{ МэВ}/c^2$ . Ими и оказались открытые в 1947 г.  $\pi$ -мезоны (см. разд. 44.2). Эти частицы нестабильны: заряженные  $\pi^\pm$ -мезоны распадаются по слабому взаимодействию (см. разд. 44.5) за время  $\tau_{\pi^\pm} = 2,6 \cdot 10^{-8}$  с. Основным типом распада является распад  $\pi^+(\pi^-)$ -мезона на мюон с зарядом того же знака и мюонное нейтрино (антинейтрино). Нейтральный  $\pi^0$ -мезон распадается на два фотона по электромагнитному взаимодействию за время порядка  $\tau_{\pi^0} = 8,4 \cdot 10^{-17}$  с. Обращаем внимание на закономерность, имеющую общий характер: чем интенсивнее взаимодействие, тем быстрее идет вызванная им реакция превращения (распада) частицы.

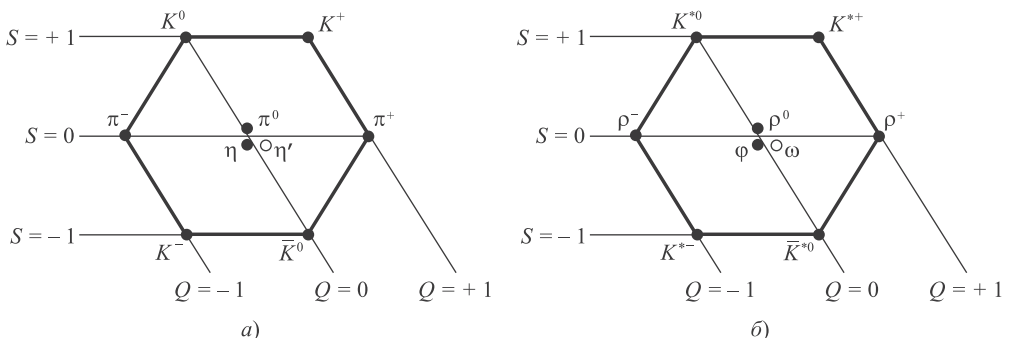
Интенсивность сильного взаимодействия очень велика — при обмене  $\pi$ -мезонами (т.е. на расстояниях порядка размера нуклона) соответствующая константа взаимодействия  $\alpha_s \sim 10\text{--}15$ , а характерное время реакций имеет порядок  $\tau_s \sim 10^{-23}$  с. Но проявляются эти силы лишь на очень малых расстояниях порядка размеров ядер.

Исследования космических лучей и эксперименты на ускорителях привели к открытию множества сильновзаимодействующих частиц — адронов. Более

того, внутри этого класса были обнаружены два больших семейства — **мезоны** (от греч. *mesos* — средний, промежуточный) с *целым спином* и **барионы** (от греч. *baros* — тяжелый), спин которых принимает *полуцелые значения*. Протон и нейтрон относятся к барионам, первым же открытым мезоном стал  $\pi$ -мезон. Было обнаружено, что при распадах и реакциях всегда сохраняется полное число барионов в системе. Поэтому барионам (в частности, нейтрону и протону) приписали **барионное число**  $B = +1$ , а для их античастиц положили  $B = -1$ . Это дало возможность провозгласить **закон сохранения барионного числа**, хотя у его физической природе мы, по сути, ничего не знаем.

Пополнение «таблицы Менделеева» для элементарных частиц производилось усилиями многих лабораторий разных стран мира. Уже в том же 1947 г. были обнаружены так называемые **странные частицы**, по всем признакам относящиеся к адронам —  $K$ -мезоны, а также лямбда-частица — барион  $\Lambda$ . Странные частицы обильно рождались в столкновениях адронов, т. е. в результате сильных ядерных взаимодействий. При этом они появлялись парами (частица с античастицей), что заставило предположить у них наличие особого квантового числа — **странности**  $S$ , закон сохранения которого объяснил бы эту особенность новых частиц. Но распадались странные частицы по слабому взаимодействию, т. е. довольно медленно, причем странность в процессах распада не сохранялась. Например, положительно заряженный  $K^+$ -мезон распадался за время  $\tau_{K^+} = 1,24 \cdot 10^{-8}$  с на лептоны или  $\pi$ -мезоны, странность которых была равна нулю.

К концу 1950-х гг., всего через 10 лет после открытия  $\pi$ -мезонов и первых странных частиц, «зоопарк» элементарных частиц вырос до трех десятков особей, и понадобилось изобрести систему, позволяющую их классифицировать. Подобная классификация была предложена в 1961 г. американским физиком М. Гелл-Манном и независимо израильским физиком и политическим деятелем Ю. Неemanом. Эту классификацию Гелл-Манн, увлекавшийся в то время буддизмом, назвал **«восьмеричным путем»** (англ. *eightfold way*), так как в ней появились восьмерки (октет) мезонов и барионов. На рисунке 45.2, *a* показано семейство мезонов со спином  $s = 1/2$ . Слева от рисунка указано значение странности  $S$  частиц, а внизу — их электрический заряд  $Q$  (в единицах элементарного заряда  $e$ ). В октет, во-первых, вошла тройка  $\pi$ -мезонов с нулевой



**Рис. 45.2.** Синглет (светлый кружок) и октет (черные кружки) мезонов со спином  $s = 0$  (*a*) и  $s = 1$  (*б*)

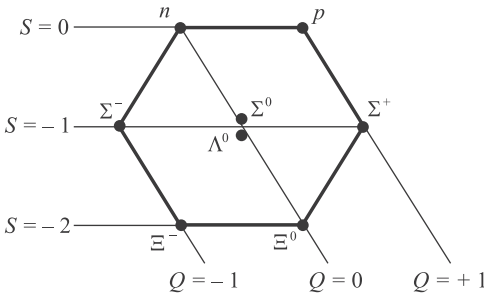


Рис. 45.3. Октет барионов со спином  $s = 1/2$

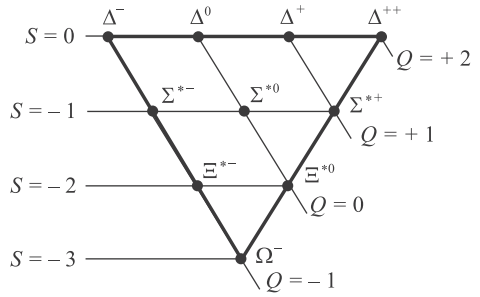


Рис. 45.4. Декуплет барионов со спином  $s = 3/2$

странностью, причем заряженные  $\pi^+$ - и  $\pi^-$ -мезоны являются античастицами друг другу, а нейтральный  $\pi^0$ -мезон является античастицей самому себе. Далее, сюда вошли заряженные  $K^+$ - и  $K^-$ -мезоны, различающиеся знаком электрического заряда и странности. Там же нашлось место и двум нейтральным странным мезонам —  $K^0$ -мезону со странностью  $S = +1$  и его античастице  $\bar{K}^0$ -мезону со странностью  $S = -1$ . Наконец, в октет вошел  $\eta$ -мезон со скрытой странностью (о чем мы поговорим в дальнейшем). Эти частицы образуют **октет псевдоскалярных мезонов**, где слово «скалярный» указывает на равенство нулю их спина, а слово «псевдо» — на отрицательную четность (см. ниже на с. 557 о свойствах квантовых систем при зеркальном отражении). Помимо этого светлым кружком на рис. 45.2, а показан примыкающий к октету **синглет** — псевдоскалярный мезон  $\eta'$  со скрытой странностью.

Существуют также векторные мезоны со спином  $s = 1$  и отрицательной четностью, обнаруженные экспериментально в начале 1960-х гг. Они входят в октет (плюс синглет), показанный на рис. 45.2, б.

Барионы также объединяются в мультиплеты. Нейтрон и протон входят в октет барионов со спином  $s = 1/2$  (рис. 45.3), а барионы со спином  $s = 3/2$  образуют декуплет (десятку) частиц (рис. 45.4). К этому времени было известно лишь девять барионов из декуплета. Гелл-Манн и Нееман предсказали существование десятого бариона  $\Omega^-$ , который вскоре был открыт экспериментаторами.

Наличие мультиплетов указывало на какую-то симметрию между странными и обычными частицами, но если бы эта симметрия была точной, совпадали бы массы частиц, входящих в один мультиплет, а они значительно отличаются друг от друга (табл. 45.1). Оказалось весьма и весьма непростым делом углядеть в ворохе экспериментальных данных эту нарушенную симметрию. В конечном итоге восьмеричный путь привел к модели **кварков**, предложенной в 1964 г. Гелл-Манном и независимо его соотечественником Дж. Цвейгом<sup>2</sup>. Согласно

<sup>2</sup>Гелл-Манн заимствовал слово «кварк» из романа Дж. Джойса «Поминки по Финнегану». Там чайки и прочие морские птицы пели песню, начинающуюся словами «Три кварка для мастера Марка» (*Three quarks for Muster Mark!*). Ничего не значащее слово (впрочем, в немецком *quark* означает творог) показалось Гелл-Манну подходящим для названия трех математических объектов, не существующих, как тогда думали, в реальности. Цвейг назвал их «тузами», но это название не прижилось.



Таблица 45.1. Свойства некоторых адронов

Мезоны	Состав	Масса (МэВ/ $c^2$ )	Барионы	Состав	Масса (МэВ/ $c^2$ )
$\pi^-$	$\bar{u}d$	139,6	$p$	$uud$	938,3
$\pi^+$	$u\bar{d}$	139,6	$n$	$udd$	939,6
$\pi^0$	$(u\bar{u} - d\bar{d})/\sqrt{2}$	135,0	$\Lambda^0$	$[ud]s$	1115,7
$K^+$	$u\bar{s}$	493,7	$\Sigma^+$	$uus$	1189,4
$K^-$	$\bar{u}s$	493,7	$\Sigma^0$	$\{ud\}s$	1192,6
$K^0$	$d\bar{s}$	497,6	$\Sigma^-$	$dds$	1197,4
$\bar{K}^0$	$\bar{d}s$	497,6	$\Delta^{++}$	$uuu$	1232
$\eta$	$(u\bar{u} + d\bar{d} - 2s\bar{s})/\sqrt{6}$	957,8	$\Xi^0$	$uss$	1314,9
$\eta'$	$(u\bar{u} + d\bar{d} + s\bar{s})/\sqrt{3}$	957,8	$\Omega^-$	$sss$	1672

этой модели, классификация адронов следует из предположения, что **мезоны построены из кварка и антикварка, а барионы — из трех кварков**. Спины кварков  $s = 1/2$ , поэтому спин мезонов должен быть равен либо нулю, либо единице (если речь идет о невозбужденных состояниях с нулевым орбитальным моментом кварков). Спин же барионов должен быть полуцелым: либо  $s = 1/2$ , либо  $s = 3/2$ . В 1960-е гг. для классификации известных к тому времени адронов было достаточно модели с тремя кварками. Их назвали  $u$ -кварк (от англ. *up* — верхний),  $d$ -кварк (от англ. *down* — нижний) и  $s$ -кварк (от англ. *strange* — странный). Необычной особенностью модели была дробность электрических зарядов кварков:  $u$ -кварк нес заряд  $2/3$  (в единицах элементарного электрического заряда),  $d$ - и  $s$ -кварки обладали зарядом  $-1/3$ . Кроме того,  $s$ -кварк нес странность  $S = -1$ , а странность  $u$ - и  $d$ -кварков равнялась нулю. Барионное число всех трех кварков равнялось  $B = 1/3$ . У антикварков  $\bar{u}$ ,  $\bar{d}$ ,  $\bar{s}$  электрический заряд, барионное число и странность имеют противоположные знаки.

Так, протон составлен из кварков  $uud$  (полный заряд  $Q = \frac{2}{3} + \frac{2}{3} - \frac{1}{3} = +1$ ), барионное число  $B = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$ , а странность  $S = 0$ . Нейтрон состоит из кварков  $udd$  (электрический заряд  $Q = \frac{2}{3} - \frac{1}{3} - \frac{1}{3} = 0$ ), барионное число  $B = 1$ , странность  $S = 0$ . Заряженный  $\pi^+$ -мезон состоит из кварка и антикварка  $u\bar{d}$  (электрический заряд  $Q = \frac{2}{3} + \frac{1}{3} = +1$ ), барионное число  $B = \frac{1}{3} - \frac{1}{3} = 0$ . Странный  $K^-$ -мезон построен из кварков  $\bar{u}s$ , так что  $Q = -\frac{2}{3} - \frac{1}{3} = -1$ , барионное число  $B = -\frac{1}{3} + \frac{1}{3} = 0$ , странность  $S = 0 - 1 = -1$ . Свойства некоторых из упомянутых частиц и их кварковый состав показаны в табл. 45.1. Заметим, что три симметричные комбинации кварка и антикварка  $u\bar{u}$ ,  $d\bar{d}$  и  $s\bar{s}$  входят в состав нейтральных мезонов  $\pi^0$ ,  $\eta$  и  $\eta'$  в виде смесей состояний. Похожим образом нейтральные барионы  $\Lambda^0$  и  $\Sigma^0$  имеют тот же кварковый состав и различаются тем, что в одном случае состояние симметрично относительно перестановки кварков  $u$  и  $d$  (что обозначено как  $\{ud\}$ ), а в другом антисимметрично (что обозначено символом коммутатора  $[ud]$ ).

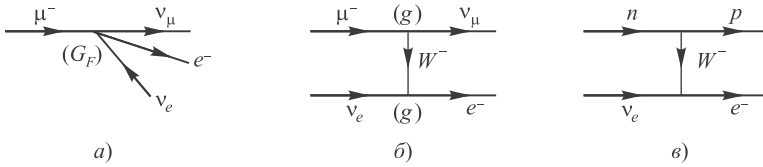
Симметрия между кварками нарушалась их массами: странный  $s$ -кварк предполагался тяжелее кварков  $u$  и  $d$ , что объясняло разницу в массах между частицами в мезонном октете, барионном октете и барионном декуплете.

Реальность кварков подтверждается в многочисленных экспериментах. Так, американские физики Дж. А. Фридман, Р. Э. Тейлор и Г. У. Кендалл, работавшие на самом мощном в то время линейном ускорителе электронов в Стэнфорде (Калифорния), провели в 1967–1973 гг. эксперименты по бомбардировке протонов электронами. Подобно Резерфорду, открывшему в 1911 г. атомное ядро, они наблюдали отскоки электронов на большие углы (так называемое *глубоко-неупругое рассеяние*), что свидетельствовало о наличии внутри протона более мелких образований. В 1990 г. эти ученые получили Нобелевскую премию «за пионерские исследования глубоконеупругого рассеяния электронов на протонах и связанных нейтронах, что имело большое значение для развития кварковой модели в физике частиц». Эти и другие эксперименты привели к тому, что к середине 1970-х гг. *кварковая модель адронов* стала доминирующей у физиков.

## Слабые взаимодействия

*Слабые взаимодействия* ответственны за  $\beta$ -распады атомных ядер (разд. 44.5), при которых излучаются электроны и антинейтрино (или позитроны и нейтрино). В частности, благодаря им нейтрон распадается на протон и электрон:  $n \rightarrow p + e^- + \bar{\nu}_e$ . Первую теорию  $\beta$ -распада создал Э. Ферми (1934). Он рассматривал процесс распада нейтрона как переход тяжелой частицы из состояния «нейтрон» в состояние «протон» с излучением двух легких частиц — электрона и антинейтрино, которых ранее в нейтроне не было. Образцом для Ферми послужила квантовая электродинамика. В 1936 г. К. Андерсон обнаружил частицу мюон с массой  $m_\mu = 207m_e$ , распадающуюся примерно за две микросекунды на электрон и два нейтрино:  $\mu^- \rightarrow e^- + \nu_\mu + \bar{\nu}_e$ . Впоследствии выяснилось, что эти нейтрино различаются и одно из них всегда появляется в паре с мюоном, а другое — в паре с электроном; поэтому их стали помечать значком соответствующей заряженной частицы. В конце 1950-х гг. теория Ферми эволюционировала в четырехфермионную теорию, в которой учитывались новые обнаруженные процессы слабого взаимодействия, но сохранялась идейная сторона ее предшественницы. В отличие от КЭД, в этой теории элементарным актом было локальное (в одной точке) взаимодействие четырех фермионов. На диаграмме Фейнмана такой процесс можно изобразить четырьмя линиями, исходящими из одной вершины. Подобно тому как в КЭД с вершиной на диаграмме Фейнмана ассоциируется электрический заряд, в теории Ферми с четырехфермионной вершиной ассоциируется параметр  $G_F = 1,44 \cdot 10^{-62} \text{ Дж} \cdot \text{м}^3 = 0,9 \cdot 10^{-52} \text{ ГэВ} \cdot \text{м}^3$ , называемый фермиевской константой связи. С ее помощью можно ввести безразмерную константу связи слабых взаимодействий  $\alpha_w$  — аналог постоянной тонкой структуры  $\alpha_{em}$  в КЭД:

$$\alpha_w = G_F \frac{(m_p c^2)^2}{(\hbar c)^3} = 1,03 \cdot 10^{-5}. \quad (45.1)$$



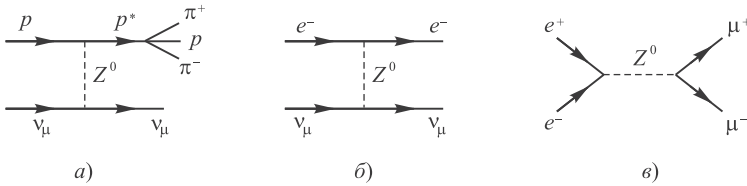
**Рис. 45.5.** Диаграмма Фейнмана для распада мюона: а) в четырехфермионной теории, б) в теории с промежуточным бозоном  $W$ ; в) аналогичная диаграмма описывает  $\beta$ -распад нейтрона

Сравнивая  $\alpha_w$  с  $\alpha_{em}$ , видим, что электромагнитное взаимодействие интенсивнее слабого почти в тысячу раз.

Существовало, однако, важное отличие четырехфермионной теории от КЭД: в последней константа связи  $\alpha_{em} = e^2/(4\pi\epsilon_0\hbar c)$  была безразмерной, а в константу Ферми  $G_F$ , как следует из (45.1), входит обратный квадрат массы. Это имело глобальные последствия для теории: несмотря на малость параметра разложения  $\alpha_w$  по теории возмущений, четырехфермионная теория слабого взаимодействия оказалась **ненормируемой** — в ней не работали рецепты КЭД, устраняющие бесконечные выражения из результатов расчетов. Стало понятно, что четырехфермионное взаимодействие на самом деле нелокально, что должны существовать переносчики слабых взаимодействий — промежуточные векторные бозоны  $W$ , а тогда диаграммы Фейнмана будут выглядеть почти так же, как в КЭД.

На рисунке 45.5, а показан процесс распада мюона на электрон и два нейтрино в четырехфермионной теории, а на рис. 45.5, б — в теории с промежуточным векторным бозоном (штриховая линия), когда мюон  $\mu^-$  сначала распадается на мюонное нейтрино  $\nu_\mu$  и виртуальный бозон  $W^-$ , а он, в свою очередь, распадается на реальные электрон  $e^-$  и электронное антинейтрино  $\bar{\nu}_e$ :  $\mu^- \rightarrow \nu_\mu + W^- \rightarrow \nu_\mu + e^- + \bar{\nu}_e$ . На этих рисунках в скобках показаны константы взаимодействия, ассоциируемые с вершинами: константа Ферми  $G_F$  в четырехфермионной теории и слабый заряд  $g$  в теории с промежуточным бозоном. Аналогичная диаграмма (рис. 45.5, в) описывает  $\beta$ -распад нейтрона. Бозон  $W^-$  уносит от мюона (нейтрона) отрицательный заряд и тем самым превращает его в мюонное нейтрино (протон) на диаграмме б (в). Далее виртуальный  $W^-$ -бозон распадается на электрон и электронное антинейтрино (линия которого направлена против стрелы времени). В такой теории должен существовать также и положительно заряженный бозон  $W^+$ , участие которого объясняет аналогичный распад положительно заряженного мюона  $\mu^+ \rightarrow \bar{\nu}_\mu + W^+ \rightarrow \bar{\nu}_\mu + e^+ + \nu_e$ . Точно такая же диаграмма (в) описывает рассеяние электронных нейтрино на нейтронах, в результате чего рождаются протон и электрон. Такое превращение наблюдалось при взаимодействии потока нейтрино с тяжелой водой  $D_2O$ , когда нейтрино превращало один из нейтронов ядра дейтерия в протон и электрон:  $\nu_e + d \rightarrow p + p + e^-$ . Во всех подобных процессах заряженный лептон превращается в соответствующее ему нейтрино (или наоборот). На диаграмме Фейнмана это изображается непрерывной линией с вершиной, в которой изменяется заряд лептона. В таком случае говорят, что вдоль линии течет заряженный слабый ток. Распад мюона изображен диаграммой, в которой имеются два слабых тока — электронный и мюонный. Поэтому говорят, что этот и аналогичные процессы вызваны взаимодействием заряженных слабых токов.

Однако наблюдались процессы, когда подобного превращения не происходило. Такие процессы (говорят, что они вызваны взаимодействием *нейтральных слабых токов*) были предсказаны теоретически и обнаружены экспериментально. Они требуют



**Рис. 45.6.** Диаграммы Фейнмана для процессов взаимодействия нейтральных токов: а) рассеяние нейтрино на нуклонах (протонах) с рождением адронов; б) рассеяние мюонных нейтрино на электронах; в) аннигиляция электрон-позитронной пары в мюонную пару (этот процесс может также идти с обменом фотоном вместо  $Z^0$ , т.е. в результате электромагнитного взаимодействия)

введения в теорию нейтральных бозонов, обозначаемых символом  $Z^0$ . Например, при рассеянии мюонных нейтрино на нуклонах заряженного мюона в продуктах реакции не наблюдалось, хотя акт взаимодействия был налицо — нуклоны переходили в возбужденное состояние (которое будем обозначать звездочкой), распадающееся затем на несколько адронов:  $\nu_\mu + p \rightarrow \nu_\mu + p^* \rightarrow \nu_\mu + p + \pi^+ + \pi^-$  (рис. 45.6, а). Обнаружились процессы рассеяния мюонных нейтрино на электронах (рис. 45.6, б) и аннигиляции электрона и позитрона в пару мюонов  $e^- + e^+ \rightarrow \mu^+ + \mu^-$  (рис. 45.6, в).

Слабое взаимодействие распространяется на очень малые расстояния порядка  $R \sim 10^{-18}$  м. Излучение виртуального переносчика слабых взаимодействий массой  $M$  создает неопределенность в энергии  $\Delta E \sim Mc^2$ . Из соотношения неопределенностей следует время жизни этой виртуальной частицы  $\Delta t \sim \hbar/\Delta E$ , за которое она может преодолеть расстояние  $R \sim c\Delta t$ . Отсюда получаем оценку  $\Delta E \sim \hbar c/R$ . В энергетических единицах  $\hbar c = 1,97 \cdot 10^{-7}$  эВ · м, откуда находим  $\Delta E \sim 100$  ГэВ. В физике элементарных частиц принято выражать массу также в энергетических единицах, так что масса промежуточной частицы должна составлять около  $100 \text{ ГэВ}/c^2$ , т.е. она на два порядка массивнее протона.

Другая оценка массы промежуточных бозонов может быть получена из сравнения диаграмм Фейнмана в теории Ферми и в теории с промежуточным векторным бозоном (см. рис. 45.5). Согласно правилам построения фейнмановских диаграмм вклад промежуточного бозона (штриховая линия на диаграмме) в процесс распада мюона обратно пропорционален квадрату энергии (или массы) бозона. Так как масса бозона велика, штриховая линия в первом приближении словно стягивается в точку, так что вклад двух вершин ( $g^2$ ) и промежуточного бозона должен дать константу Ферми:  $G_F \propto g^2/M^2$ . Если предположить, что слабый заряд  $g$ , ассоциированный с вершинами в теории с промежуточными бозонами, примерно равен электрическому заряду, то соотношение (45.1) для константы связи слабого взаимодействия может быть записано в виде

$$\alpha_w \sim \alpha_{em} \frac{m_p^2}{M^2} = 1,03 \cdot 10^{-5}, \tag{45.2}$$

откуда следует оценка массы промежуточного бозона

$$M \sim m_p \sqrt{\frac{\alpha_{em}}{\alpha_w}} \sim 27m_p \approx 25 \text{ ГэВ}/c^2. \tag{45.3}$$

Конечно, нельзя слишком сильно доверять полученным числам — это всего лишь оценки по порядку величины. Но они позволили получить представление, в какой области энергий можно искать промежуточные векторные бозоны.

Теория с тремя промежуточными бозонами  $W^+$ ,  $W^-$  и  $Z^0$  была развита американцами Ш. Глэшоу, С. Вайнбергом и пакистанским физиком А. Саламом. В основу была положена симметрия слабых взаимодействий относительно замены протонов нейтронами и наоборот, а также взаимозаменяемость нейтрино и электронов. Первые открытые кварки и лептоны естественным образом разбили на пары, которые можно представить столбцами:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}. \quad (45.4)$$

Естественно, есть еще две пары соответствующих античастиц — антикварки  $\bar{u}$ ,  $\bar{d}$  и антинейтрино  $\bar{\nu}_e$  с позитроном  $e^+$ . В дальнейшем мы без особой необходимости не будем упоминать об антикварках и антилептонах, помня, однако, что они существуют. Частицы в парах расположены так, что электрические заряды частиц в верхних строках столбцов в (45.4) на единицу элементарного заряда больше зарядов частиц в нижних строках (разумеется, для соответствующих столбцов из античастиц заряды «верхних» частиц на единицу *меньше* зарядов «нижних»).

Различия в зарядах частиц в верхних и нижних строках касаются их электромагнитных взаимодействий. Слабые же взаимодействия «не видят» разницы между кварками  $u$  и  $d$  и лептонами  $\nu_e$  и  $e^-$ . Поэтому вместо этих пар можно использовать другие, являющиеся их суперпозицией. Операторы, «перемешивающие» кварки и лептоны внутри пар, действуют на столбцы (45.4) и представляются, очевидно, матрицами  $2 \times 2$ . Набор таких матриц образует, как говорят, группу  $SU(2)$ . Произведение двух матриц некоммутативно, т. е. результат зависит, вообще говоря, от порядка их действия — о таких группах говорят, что они *неабелевы*. Если коэффициенты этих матриц не зависят от координат, мы имеем дело с *глобальной калибровочной инвариантностью* теории слабых взаимодействий. Требование *локальной калибровочной инвариантности* относительно группы  $SU(2)$ , когда коэффициенты матриц зависят от радиуса-вектора, вместе с абелевой локальной симметрией КЭД (ее обозначают  $U(1)$ ) привело к полям Янга–Миллса, кванты которых — фотон  $\gamma$  и промежуточные бозоны  $W^+$ ,  $W^-$  и  $Z^0$  — имеют нулевую массу. Если бы симметрия между электромагнетизмом и слабым взаимодействием была точной, промежуточные бозоны были бы родными братьями фотона. Но симметрия оказывается нарушенной из-за взаимодействия калибровочных бозонов со скалярным полем Хиггса (в современной терминологии это называется *механизмом Браута–Энглера–Хиггса*), благодаря чему промежуточные бозоны приобретают массу, а фотон остается безмассовым. В 1979 г. Глэшоу, Салам и Вайнберг получили Нобелевскую премию «за вклад в объединенную теорию слабых и электромагнитных взаимодействий между элементарными частицами, в том числе предсказание слабых нейтральных токов» (см. с. 554).

Эта теория называется *теорией электрослабых взаимодействий*. Из нее следует, что электромагнитное и слабое взаимодействия различаются при низких энергиях, но при энергиях более 100 ГэВ уже неотличимы друг от друга.

Результаты теории электрослабых взаимодействий, касающиеся масс промежуточных векторных бозонов, оказались близки к оценке (45.3), отличаясь от нее лишь численными множителями и зависимостью от так называемого угла

Вайнберга  $\theta_W$ , связывающего слабый заряд с электрическим ( $g = e/\sin\theta_W$ ):

$$M_W = \frac{37,28}{\sin\theta_W} \text{ ГэВ}/c^2, \quad M_Z = \frac{M_W}{\cos\theta_W}. \quad (45.5)$$

Значение угла Вайнберга теория не предсказывает, но оно было найдено по данным экспериментов еще до экспериментального открытия промежуточных бозонов, откуда следовали предсказания теорией масс бозонов:  $M_W \sim 80 \text{ ГэВ}/c^2$  и  $M_Z = 90 \text{ ГэВ}/c^2$ . Промежуточные векторные бозоны были открыты в 1983 г. на протон-антипротонном коллайдере в Европейской организации ядерных исследований (ЦЕРН, Женева). Массы промежуточных векторных бозонов оказались в предсказанном теоретиками диапазоне:  $M_W = 80,385 \pm 0,015 \text{ ГэВ}/c^2$ ,  $M_Z = 91,1876 \pm 0,0021 \text{ ГэВ}/c^2$ . Уже в 1984 г. итальянец К. Руббиа и голландец С. ван дер Мейр получили Нобелевскую премию «за решающий вклад в большой проект, осуществление которого привело к открытию квантов поля  $W$  и  $Z$  — переносчиков слабого взаимодействия».

Объединенная теория электрослабых взаимодействий принесла еще ряд Нобелевских премий, в частности в 1999 г. голландцам М. Вельтману и Г. т'Хоофту за доказательство ее *перенормируемости*, т. е. возможности устранения возникающих при вычислениях бесконечных выражений. По выражению одного из известных физиков, это «превратило лягушку Вайнберга—Салама в очаровательную принцессу». Теория электрослабых взаимодействий неизменно подтверждалась в экспериментах. Но долгое время оставался неуловимым **бозон Хиггса** — квант поля, взаимодействие с которым придает массу промежуточным векторным бозонам  $W^\pm$  и  $Z_0$ . Его существование было постулировано теоретиками еще в 1964 г., но никаких предсказаний о его массе теория не давала. Только в 2012 г. бозон Хиггса с массой  $m_H = 125,7 \pm 0,4 \text{ ГэВ}/c^2$  был обнаружен экспериментально. В 2013 г. англичанин П. Хиггс и бельгиец Ф. Энглер были удостоены Нобелевской премии «за теоретическое обнаружение механизма, который помогает нам понять происхождение масс субатомных частиц, подтвержденного в последнее время обнаружением предсказанной элементарной частицы в экспериментах *ATLAS* и *CMS* на Большом адронном коллайдере в ЦЕРН». Открытие бозона Хиггса завершило создание теории электрослабых взаимодействий.

## Нарушение $P$ -, $C$ - и $CP$ -симметрии в слабых взаимодействиях

Прежде чем перейти к другим фундаментальным взаимодействиям, мы поговорим о *дискретных симметриях* — инвариантности законов природы относительно различных операций отражения.

При операции  $\hat{P}$  (от англ. *parity* — четность) все три координатные оси меняют направление на противоположное. Легко сообразить, что при обычном отражении в зеркале направление меняет лишь одна ось, ортогональная его плоскости. Но если вдобавок повернуть зеркальное отражение вокруг этой оси на  $180^\circ$ , мы получим операцию  $\hat{P}$ . Живая природа не обладает зеркальной симметрией: достаточно вспомнить, что сердце у подавляющего большинства людей находится слева, а печень — справа или что на Земле 97 % правшей и только 3 % левшей. Зеркальная симметрия нарушается уже на молекулярно-химическом уровне: преимущественное направление закрутки наблюдается у спиралей молекул аминокислот, белков, нуклеиновых кислот,

сахаров и т. д. Объяснений этим фактам пока не найдено. Но в простейших системах, изучаемых физикой, такого различия не наблюдается. Засняв, например, на видео отражение в зеркале движения тела, брошенного под углом к горизонту, не удастся потом обнаружить расхождения с законами механики. Долгое время предполагали, что законы физики инварианты относительно операции  $\hat{P}$ , чему соответствовал закон сохранения *пространственной четности*.

Частица, рассматриваемая в зеркале, должна оставаться той же частицей. Это значит, что ее волновая функция при операции отражения может приобрести лишь несущественный множитель:  $\hat{P}\psi = \eta\psi$ . Повторим операцию отражения:  $\hat{P}^2\psi = \eta\hat{P}\psi = \eta^2\psi$ . Но дважды повторенная операция отражения возвращает систему в первоначальное состояние:  $\hat{P}^2\psi = \psi$ . Отсюда следует, что каждой частице можно приписать определенную *внутреннюю четность* — положительную ( $\eta = +1$ ) или отрицательную ( $\eta = -1$ ).

В 1956 г. китайско-американские физики Ли Цзундао и Янг Чжэньнин предсказали нарушение четности в процессах слабого взаимодействия, а Ву Цзяньсюн обнаружила его экспериментально. В 1957 г. Ли и Янг получили Нобелевскую премию «за проницательное исследование так называемых законов четности, которое привело к важным открытиям в физике элементарных частиц». В опыте Ву измерялась интенсивность потока электронов, испущенных при  $\beta$ -распаде ядер кобальта  $^{60}\text{Co} \rightarrow ^{60}\text{Ni} + e^- + \bar{\nu}_e$ , но обсуждаемое явление проще понять на примере распада пи-мезонов на мюоны и нейтрино:

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu, \quad (45.6)$$

$$\pi^+ \rightarrow \mu^+ + \nu_\mu. \quad (45.7)$$

Эксперименты показывают, что в реакции (45.6) спин мюона всегда параллелен его импульсу, а в реакции (45.7) спин мюона всегда антипараллелен его импульсу. При отражении в зеркале все обычные (полярные) векторы (в частности, импульс) меняют направление, а аксиальные векторы (угловой момент или спин) — нет. Поэтому при отражении реакции (45.6) в зеркале мы увидим распад  $\pi^-$ -мезона, при котором спин и импульс рождающегося мюона антипараллельны. Такого процесса нет в природе, это и значит, что пространственная четность не сохраняется. Аналогичный вывод можно сделать при отражении в зеркале реакции (45.7).

Причина нарушения зеркальной симметрии кроется в свойствах нейтрино. Масса нейтрино очень мала, и оно движется практически со скоростью света, причем спин нейтрино направлен против его импульса. Говорят, что нейтрино — частица с *левой (отрицательной) спиральностью* (напоминает винт с левой резьбой). Антинейтрино — частица с *правой (положительной) спиральностью*, его спин параллелен импульсу. Теперь ясно, что зеркальным изображением процесса с испусканием нейтрино будет процесс с участием нейтрино с правой спиральностью, но до сих пор такой частицы в природе не наблюдалось.

Операция *зарядового сопряжения*  $\hat{C}$  (от англ. *charge* — заряд) меняет знаки всех зарядов частицы (не только электрического, но и лептонного числа или любого другого заряда, какой еще у частицы найдется). Иными словами, операция  $\hat{C}$  превращает частицу в античастицу:  $\pi^+$ -мезон превратится в  $\pi^-$ -мезон, отрицательно заряженный мюон — в положительно заряженный мюон, антинейтрино — в нейтрино и т. п. Применение операции  $\hat{C}$  к обеим частям реакции (45.6) превращает ее в реакцию (45.7), но направления спина и импульса мюона остаются по-прежнему параллельными. Таким образом, снова получается несуществующий в природе процесс; *зарядовая четность* тоже не сохраняется в слабых взаимодействиях, хотя античастицы ничем существенным от частиц не отличаются. Определенную зарядовую четность можно приписать лишь истинно нейтральным частицам, которые совпадают со своими античастицами (фотону,

$\pi^0$ -мезону и т. п.). Волновые функции таких частиц приобретают множители  $+1$  или  $-1$  (рассуждения здесь точно такие же, как и в случае пространственной четности). В процессах сильного и электромагнитного взаимодействий зарядовая четность сохраняется, что приводит к определенным ограничениям на реакции распада элементарных частиц.

Открытие нарушения  $P$ -симметрии произвело большое впечатление на физиков, поскольку означало асимметрию самой Вселенной по отношению к зеркальному отражению. Казалось, душевное равновесие встревоженных ученых было восстановлено с помощью операции **комбинированной инверсии**  $\hat{C}\hat{P}$ , при которой одновременно заменяется нейтрино на антинейтрино и левая спиральность на правую. В соответствии с этой гипотезой, предложенной Л. Д. Ландау (1957), мир, потеряв  $P$ - и  $C$ -симметрию, оставался бы  $CP$ -симметричным, в нем сохранялась бы **комбинированная  $CP$ -четность**. Если бы все взаимодействия были инвариантными относительно операции комбинированной инверсии, пространство само по себе оставалось бы симметричным при отражении, асимметричными были бы лишь заряды частиц.

Однако в 1964 г. американские физики Дж. Кронин и В. Фитч обнаружили в опытах по распаду нейтральных  $K^0$ -мезонов (каонов) нарушение  $CP$ -симметрии. Эти частицы обладают удивительным свойством:  $K^0$  могут превращаться в свои античастицы  $\bar{K}^0$  и наоборот (такое стало возможно из-за того, что  $K^0$  и  $\bar{K}^0$  различаются лишь странностью, которая в слабых взаимодействиях не сохраняется). Из-за этого явления, называемого **осцилляциями нейтральных  $K$ -мезонов**, физически наблюдаемыми являются суперпозиции нейтральных  $K^0$  и  $\bar{K}^0$  мезонов, массы которых слегка различаются. Предполагалось, что коэффициенты в этих суперпозициях одинаковы по абсолютной величине и, стало быть, соответствующие частицы имеют определенную  $CP$ -четность: так называемый долгоживущий  $K_L$  мезон с  $CP=-1$  и временем жизни  $5,2 \cdot 10^{-8}$  с представлял собой разность состояний с противоположной странностью, а короткоживущий  $K_S$  мезон с  $CP=+1$  и временем жизни  $0,9 \cdot 10^{-10}$  с был их суммой:

$$K_L = \frac{K^0 - \bar{K}^0}{\sqrt{2}}, \quad K_S = \frac{K^0 + \bar{K}^0}{\sqrt{2}}. \quad (45.8)$$

В распадах на  $\pi$ -мезоны при условии сохранения  $CP$ -четности долгоживущий каон должен был распадаться на три пи-мезона ( $K_L \rightarrow \pi^+ + \pi^- + \pi^0$  или  $3\pi^0$  — эти системы имеют  $CP=-1$ ), а короткоживущий каон обязан был распадаться на два пи-мезона ( $K_S \rightarrow \pi^+ + \pi^-$  или  $2\pi^0$  —  $CP=+1$  для этих систем). Однако Кронин и Фитч наблюдали изредка распады долгоживущих  $K_L$ -мезонов на два  $\pi$ -мезона. Это означало, что абсолютные значения коэффициентов в суперпозициях (45.8) слегка различаются, а  $CP$ -четность не сохраняется. В 1980 г. Кронин и Фитч были удостоены Нобелевской премии «за открытие нарушений фундаментальных принципов симметрии в распаде нейтральных  $K$ -мезонов».

## 45.2 Квантовая хромодинамика

Когда число известных адронов перевалило за триста, стало ясно, что их слишком много, чтобы они могли претендовать на роль фундаментальных составляющих вещества и быть полноправными членами аналога таблицы Менделеева для микромира. (Однако по инерции их продолжают называть элементарными частицами — видимо, не придумали более удачного названия.) Поэтому физики задумались над моделями, позволяющими построить известные науке адроны из более фундаментальных составляющих. Хорошими кандидатами на такую роль оказались **кварки** (см. с. 551), которые поначалу рассматривались лишь как



элементы удачной математической конструкции, позволившей провести инвентаризацию адронов и распределить их по октетам и декуплетам, подобно тому как Карл Линней распределил живые организмы и растения по видам, родам, отрядам и классам. О свойствах кварков мы и поговорим в этом разделе.

Вскоре после рождения кварковой модели адронов выяснилось, что помимо спина, электрического заряда, барионного числа и странности у кварков должно быть еще какое-то свойство (квантовое число), отличающее их друг от друга. Более того, это число должно принимать три значения. Действительно, согласно модели кварков открытый экспериментально в 1964 г. барион  $\Omega^-$  должен состоять из трех кварков  $sss$ . Спин этой частицы  $s = 3/2$ , т.е. все три одинаковых странных кварка находятся внутри нее в одинаковом состоянии. Но это противоречит принципу Паули. Аналогичная проблема сопровождала открытый еще в 1952 г.  $\Delta^{++}$ -барион, тоже состоящий из трех одинаковых кварков  $uuu$  с параллельными спинами.

Выход был найден в предположении о наличии у кварков ранее неизвестного квантового числа<sup>3</sup>. Для объяснения существования барионов  $\Omega^-$  и  $\Delta^{++}$  дополнительное квантовое число должно было принимать три значения. Таким образом, число состояний кварков утроилось, а число адронов осталось прежним, не зависящим от упражнений теоретиков. Отсюда следовало, что не все комбинации кварковых состояний приводят к новым наблюдаемым частицам, что тоже требовало объяснения.

Были предложены разные схемы реализации идеи о дополнительном квантовом числе (в том числе модели с целочисленными электрическими зарядами кварков), но в конечном итоге общепринятой и подтвержденной в экспериментах схемой стала теория М. Гелл-Манна и Х. Фрича (1971). В этой теории электрические заряды кварков были дробными, как в старой кварковой модели, а новое квантовое число авторы называли *цветом*. Это квантовое число «окрашивает» кварк в один из трех цветов: красный, зеленый или синий (названия, очевидно, условны, но образны). Соответственно антикварки должны быть «окрашены» в антицвета: антикрасный, антизеленый и антисиний. Следует подчеркнуть, что цвет кварков ничего общего не имеет с обычным цветом, где любые оттенки могут быть получены смешиванием трех базовых цветов, — название возникло всего лишь по аналогии.

Наблюдаемые адроны (мезоны и барионы) не «окрашены», они составлены из «бесцветных» комбинаций кварков. Это объясняет, почему утроение числа цветных кварковых состояний не приводит к увеличению числа наблюдаемых адронов. Отсюда следуют свойства цвета.

---

<sup>3</sup>Честь первого упоминания необходимости дополнительного квантового числа у кварков принадлежит ученым Объединенного института ядерных исследований (ОИЯИ, Дубна): сначала Б. В. Струминский, в то время аспирант Н. Н. Боголюбова, написал в своей работе (7.01.1965), что кваркам «надо приписать новое квантовое число», а чуть позже, но в том же месяце в более подробной статье эту идею разработали Н. Н. Боголюбов, Б. В. Струминский и А. Н. Тавхелидзе. В том же 1965 году независимо от дубненцев такую же мысль высказали американские физики М.-Ё. Хан и Й. Намбу, японский физик И. Миямото, а также американский физик П. Фройнд. Н. Н. Боголюбов отмечал, что Хан и Намбу сделали «принципиально новый шаг на пути развития динамической теории адронов», впервые введя в рассмотрение векторные поля — переносчики сильного взаимодействия.

- Любой цвет компенсируется своим антицветом. Это объясняет существование наблюдаемых бесцветных мезонов, состоящих из цветного кварка и антикварка, несущего антицвет. Скажем, мезон  $\pi^+$  состоит из красного (зеленого, синего) кварка  $u$  и антикрасного (антизеленого, антисинего) антикварка  $\bar{d}$  и потому бесцветен.
- Смесь всех трех цветов дает бесцветное («белое») состояние. Это объясняет бесцветность барионов, состоящих из трех кварков. Можно взять, например, красный кварк  $u$ , зеленый кварк  $u$  и синий кварк  $d$ . Их связанное состояние (протон) в целом нейтрально по отношению к цветовому заряду (т. е. бесцветно).

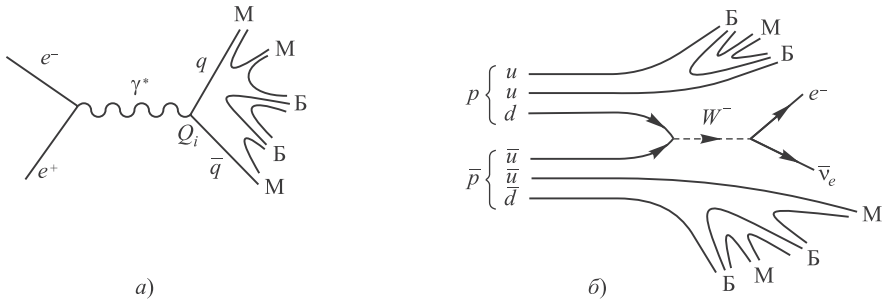
В 1972 г. в работе Гелл-Манна и Фрича и в их совместной работе с швейцарским теоретиком Г. Лейтвилером была создана квантовая теория цветовых взаимодействий трех цветных кварков с дробными электрическими зарядами — **квантовая хромодинамика (КХД)**. Согласно КХД, цвет — это уже не просто новое квантовое число наподобие барионного или лептонного чисел, а **цветовой заряд**, порождающий особое поле, подобно тому как заряд электрический порождает электрическое поле. Как и теория электрослабых взаимодействий, КХД опиралась на теорию калибровочных полей Янга—Миллса. Предполагалось, что цветовая симметрия является точной, так что кварки данного сорта ( $u$ , либо  $d$ , либо  $s$ ) имеют один и тот же электрический заряд, одинаковые барионное число и странность, различаясь только цветом. Три цветовых состояния каждого кварка можно представить столбцом с тремя строками:

$$u_{\text{к}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad u_{\text{з}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad u_{\text{с}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (45.9)$$

Соответственно оператор, перемешивающий цветовые состояния, должен являться матрицей  $3 \times 3$ . Требование **локальной калибровочной цветовой симметрии SU(3)** привело к появлению в теории переносчиков цветовых взаимодействий — безмассовых бозонов со спином  $s = 1$ . Гелл-Манн и Фрич назвали их **глюонами** (от англ. *glue* — клей), так как они «склеивают» кварки в адронах.

Сильное ядерное взаимодействие перестало быть фундаментальным взаимодействием, превратившись в «остаток», «хвост» более фундаментального цветового взаимодействия между кварками. Таким образом, с сильным ядерным взаимодействием произошло примерно то же самое, что и с межмолекулярными силами Ван-дер-Ваальса (см. т. 1, разд. 11.5) — «хвостом» электромагнитных взаимодействий.

Поиски свободных кварков результатов не дали. В результате физики пришли к выводу, что частицы с «открытым цветом» в свободном состоянии не наблюдаются — взаимодействие между ними не ослабевает, а растет с расстоянием, и становится столь сильным, что вырвать кварк или глюон из адрона невозможно. Мезон, например, можно представить себе как жесткую струну, соединяющую цветные кварк и антикварк. При попытке растянуть кварки подальше друг от друга струна в какой-то момент лопается, но в месте разрыва образуется новая



**Рис. 45.7.** Диаграмма Фейнмана для амплитуды аннигиляции электрона и позитрона в адроны (*a*) и для рождения пары «электрон и нейтрино» в протон-антипротонных столкновениях (*б*)

кварк—антикварковая пара, так что мы получаем две струны с цветными объектами на концах — два обычных бесцветных мезона. Явление невылетаия кварков из адронов называется **конфайнментом**<sup>4</sup>. Гипотеза конфайнмента цветных кварков подтверждается некоторыми численными расчетами, но математически пока не доказана.

Эксперименты свидетельствовали в пользу модели цветных кварков с дробными электрическими зарядами. Так, рассматривался процесс аннигиляции электронов и позитронов в адроны (рис. 45.7, *a*). На первом этапе процесс идет с участием электромагнитных взаимодействий: при аннигиляции рождается виртуальный фотон  $\gamma^*$ , который распадается на кварк-антикварковую пару. Далее эти кварки удаляются друг от друга, но, не будучи в состоянии улететь на бесконечность из-за конфайнмента, порождают другие кварк-антикварковые пары, соединяющиеся в наблюдаемые «бесцветные» адроны — мезоны М и барионы Б. Заряд  $Q_i$  кварков ассоциируется с вершиной их взаимодействия с виртуальным фотоном, поэтому сечение (вероятность) данного  $i$ -го канала аннигиляции пропорционально  $Q_i^2$ . Суммируя по всем каналам, получаем, что полное сечение этого процесса пропорционально сумме квадратов зарядов всех кварков, способных родиться при данной инвариантной массе виртуального фотона (т.е. при данной энергии столкновения электрона и позитрона — см. т. 1, разд. 9.15). При делении на сечение реакции аннигиляции электрона и позитрона в пару мюонов сокращаются все дополнительные множители, и мы получаем для отношения этих сечений выражение

$$R = \frac{e^- + e^+ \rightarrow \text{адроны}}{e^- + e^+ \rightarrow \mu^- + \mu^+} = \sum Q_i^2. \quad (45.10)$$

Теперь для трех цветных кварков с дробными зарядами находим:

$$R = 3 \left( \frac{2^2}{3^2} + \frac{1^2}{3^2} + \frac{1^2}{3^2} \right) = 2.$$

<sup>4</sup>От англ. глагола *to confine* — ограничивать, держать взаперти, заключать в тюрьму.

Множитель 3 учитывает вклад трех цветных состояний кварков; без цвета мы имели бы  $R = 2/3$ . Эксперименты тех лет подтверждали КХД и давали при энергиях столкновения 2–10 ГэВ примерно постоянное значение  $R = 2,5 \pm 0,4$ .

Уже отмечалось, что эксперименты по глубоконеупругому рассеянию указывали на наличие внутри адронов точечных составляющих. В столкновениях  $e^-p$ ,  $pp$  и  $p\bar{p}$  в экспериментах на ускорителях кварк внутри протона или антипротона мог получить большой импульс и попытаться вылететь наружу. Растущие с расстоянием цветовые силы не позволяют ему это сделать, и такой кварк порождает *струи обычных бесцветных адронов*, экспериментальное обнаружение которых также свидетельствует в пользу гипотезы теории кварков. На рисунке 45.7, б показан процесс рождения лептонной пары в протон-антипротонных столкновениях, сопровождаемый множественным рождением адронных струй — мезонов  $M$  и барионов  $B$ , но свободные кварки в продуктах реакции не появляются.

Роль глюонов сводится к «перекрашиванию» кварка. Глюон несет пару цветов (например, синий и антикрасный) и, будучи поглощен красным кварком, компенсирует красный цвет и окрашивает кварк в синий. Нетрудно понять, что раз у нас всего три цвета, то можно составить  $3 \times 3 = 9$  цветовых комбинаций для глюонов. Но одна из них полностью нейтральна (симметричная смесь в равных долях красного и антикрасного, зеленого и антитезеленого, синего и антисинего цветов  $k\bar{k} + z\bar{z} + c\bar{c}$ ). Такого глюона в природе нет, так как он не вносит вклада в цветовое взаимодействие, потому что с равной вероятностью компенсировал бы любой из трех цветов и окрасил бы кварк снова в тот же цвет. Две другие комбинации цвета и того же антицвета ( $k\bar{k} - z\bar{z}$  и  $k\bar{k} + z\bar{z} - 2c\bar{c}$ ) вносят вклад в цветовое взаимодействие наряду с шестью комбинациями кварка и антикварка разных цветов ( $k\bar{z}$ ,  $z\bar{c}$  и т. п.). Таким образом, в квантовой хромодинамике взаимодействие между кварками осуществляется путем обмена восьмеркой глюонов.

Цветовая симметрия описывается некоммутирующими матрицами. Вследствие этого глюоны сами имеют цветовые заряды и могут взаимодействовать друг с другом. Поэтому КХД — нелинейная теория, гораздо более трудная для изучения и применения, чем КЭД, где фотон электрическим зарядом не обладает.

### 45.3 Кварки и лептоны в Стандартной теории

Когда в 1974 г. был обнаружен четвертый  $c$ -кварк (от англ. *charm* — шарм, очарование, а также магическое заклинание) с электрическим зарядом  $+2/3$ , российские физики любили цитировать в своих докладах последнюю строфу стихотворения Осипа Мандельштама «За Паганини длиннопалым» («Скрипачка»):

Играй же на разрыв аорты  
С кошачьей головой во рту,  
Три черта было — ты четвертый,  
Последний чудный черт в цвету.

Физики углядели здесь и ранее найденные три «черта» (кварка), и их цвет, и чудо (чары, очарованный). Новый очарованный кварк стал «парой» странного

кварка, завершив второе поколение кварков и лептонов (мюон и мюонное нейтрино). Он был обнаружен в составе  $J/\psi$  частиц с массой  $3,1 \text{ ГэВ}/c^2$  (в три раза тяжелее протона). Двойное название связано с тем, что эти частицы, состоящие из кварка и антикварка ( $c\bar{c}$ ), были одновременно открыты в США двумя группами экспериментаторов, работавших в Стэнфорде и Брукхейвенской национальной лаборатории (США), и получили в одном месте имя  $J$ , а в другом —  $\psi$ . Руководители этих групп Б. Рихтер и С. Тинг стали в 1976 г. лауреатами Нобелевской премии «за основополагающий вклад в открытие тяжелой элементарной частицы нового типа». «Очарование» в  $J/\psi$ -мезонах скрыто, но позже были найдены мезоны с открытым очарованием: заряженные  $D^+$  ( $c\bar{d}$ ),  $D^-$  ( $\bar{c}d$ ), нейтральные  $D^0$ ,  $\bar{D}^0$  ( $c\bar{u}$ ,  $\bar{c}u$ ), «странно-очарованные»  $D_S^+$  ( $c\bar{s}$ ) и  $D_S^-$  ( $\bar{c}s$ ). Обнаружились и очарованные барионы:  $\Lambda_C^+$  ( $[ud]c$ ),  $\Sigma_C^{++}$  ( $uuc$ ),  $\Sigma_C^+$  ( $\{ud\}c$ ),  $\Sigma_C^0$  ( $dcc$ ).

Однако, вопреки предсказанию поэта, четвертый «чудный черт в цвету» оказался не последним. Еще в 1973 г. японские теоретики М. Кобаяси и Т. Маскава обнаружили, что для описания явления нарушения  $CP$ -четности необходимо три поколения кварков. Через четыре года, в 1977 г., экспериментаторы открыли пятый кварк  $b$  (от англ. *beauty* — красота, прелесть) с электрическим зарядом, равным  $-1/3$ . «Прелестный кварк» нашелся в составе  $\Upsilon$ -мезонов ( $b\bar{b}$ ). Позже были найдены частицы с «обнаженной прелестью»: нейтральные мезоны  $B^0$  ( $d\bar{b}$ ) и  $\bar{B}^0$  ( $b\bar{d}$ ), «странно-прелестные» нейтральные мезоны  $B_s^0$  ( $s\bar{b}$ ) и  $\bar{B}_s^0$  ( $b\bar{s}$ ), «очарованно-прелестные» мезоны  $B_c^+$  ( $c\bar{b}$ ) и  $\bar{B}_c^0$  ( $b\bar{c}$ ). Поскольку  $b$ -кварк можно считать кузеном странного  $s$ -кварка (в своих парах оба являются «нижними» кварками), ожидалось и действительно наблюдались (2001) осцилляции нейтральных прелестных  $B$ -мезонов и нарушение  $CP$ -четности, подобные аналогичным явлениям в системе нейтральных странных  $K$ -мезонов.

Было ясно, что для полного состава третьего поколения кварков недостает еще одного кварка с зарядом  $2/3$ , который заранее окрестили  $t$ -кварком (от англ. *truth* — правда, истина). Он был обнаружен в 1995 г. в экспериментах по столкновению протон-антипротонных пучков на ускорителе Теватрон в Национальной лаборатории им. Э. Ферми в Батавии вблизи Чикаго (шт. Иллинойс, США). Этот кварк очень тяжел (масса около  $173 \text{ ГэВ}/c^2$ ) и живет крайне недолго ( $5 \cdot 10^{-25} \text{ с}$ ), так что в составе адронов не существует.

Еще раньше, в 1975 г. в Стэнфордском университете, были найдены лептоны третьего поколения:  $\tau$ -лептон с массой  $1778,8 \text{ МэВ}/c^2$  и временем жизни  $2,9 \cdot 10^{-13} \text{ с}$  (руководитель эксперимента М. Перл получил Нобелевскую премию в 1995 г.), а в 2000 г. в Лаборатории им. Э. Ферми в прямых экспериментах были обнаружены  $\tau$ -нейтрино.

Символы  $u, d, c, s, t, b$ , различающие сорта кварков, распалившиеся физики называли **ароматами** (раз уж в сложной квантовой теории появился цвет, то почему бы не быть и запаху?).

Подведем итоги. К настоящему времени обнаружено три дублета лептонов и — в составе адронов — три дублета кварков. Эти дублеты фундаментальных частиц называют **поколениями**. Каждой паре кварков в поколении должна соответствовать пара лептонов. Число пар лептонов и кварков должно совпадать, иначе окажется противоречивой теория, объединяющая электромагнитные и

кварки	2,3 МэВ 2/3 1/2	1,27 ГэВ 2/3 1/2	173 ГэВ 2/3 1/2	калибровочные бозоны	0 0 1	125,7 ГэВ 0 0	
	<b><i>u</i></b>	<b><i>c</i></b>	<b><i>t</i></b>		<b><i>g</i></b>		<b><i>H</i><sup>0</sup></b>
	4,8 МэВ -1/3 1/2	95 МэВ -1/3 1/2	4,2 ГэВ -1/3 1/2		<b><i>γ</i></b>		бозон Хиггса
лептоны	< 2,2 эВ 0 1/2	< 0,19 МэВ 0 1/2	< 18,2 МэВ 0 1/2	91,2 ГэВ 0 1	80,4 ГэВ ±1 1		
	<b><i>ν<sub>e</sub></i></b>	<b><i>ν<sub>μ</sub></i></b>	<b><i>ν<sub>τ</sub></i></b>	<b><i>Z</i><sup>0</sup></b>			
	0,511 МэВ -1 1/2	105,7 МэВ -1 1/2	1,777 ГэВ -1 1/2	<b><i>W</i><sup>±</sup></b>			
	I	II	III				
	поколение фермионов						

**Рис. 45.8.** Фундаментальные частицы Стандартной теории: слева от символа частицы показаны (сверху вниз) ее масса (в энергетических единицах), электрический заряд (в единицах  $e$ ) и спин; у каждого фермиона имеется также античастица с противоположным электрическим зарядом

слабые взаимодействия. Таким образом, современная «таблица Менделеева» для **фундаментальных частиц** имеет простой вид (рис. 45.8).

Однако теория не дает ответов на вопросы о числе цветов и числе поколений (три — необходимый минимум, природа обычно экономна, но кто знает). Ответы на оба вопроса ищут экспериментаторы. Поведение сечения аннигиляции электронов и позитронов в адроны в области сравнительно невысоких энергий зависит, как уже говорилось, от числа цветов  $N_c$ . Объединенные данные измерений независимых групп на разных коллайдерах дали для числа цветов результат  $N_c = 3,06 \pm 0,10$ .

Сечение этого же процесса при высоких энергиях вблизи порога рождения  $Z^0$ -бозонов зависит от числа  $N_\nu$  легких (с массой, меньшей половины массы  $Z^0$ -бозона) нейтрино. А это то же самое, что и число поколений. Объединенные результаты различных групп, приведенные в современных таблицах свойств элементарных частиц, выглядят впечатляюще:  $N_\nu = 2,984 \pm 0,008$ .

Таким образом, физики уже знают все фундаментальные частицы, из которых построено вещество: *шесть лептонов, шесть кварков и двенадцать их античастиц*. Лептоны и кварки в настоящее время рассматриваются как точечные частицы без внутренней структуры. Электрослабые взаимодействия между ними переносятся *фотоном и тремя промежуточными векторными бозонами  $W^\pm$  и  $Z^0$* , сильные цветовые взаимодействия — *глюонами, заключенными, как и кварки, внутри адронов*. Все переносчики фундаментальных взаимодействий являются **калибровочными бозонами** — квантами калибровочных полей со спином, равным единице.

Из частиц, предсказанных Стандартной теорией, последним был открыт (2012) скалярный (с нулевым спином) **бозон Хиггса**. Масштаб этих исследований потрясает. Бозон Хиггса был найден в ЦЕРНе на двух детекторах — *ATLAS*

и *CMS*. Длина детектора *ATLAS* — 45 м, диаметр — 25 м, масса — 7000 т. В создании детектора и проведении экспериментов участвовало около 3000 физиков из 175 научных организаций 38 стран. Длина детектора *CMS* — 25 м, диаметр — 15 м, масса — 12 500 т. В его создании и эксплуатации приняло участие 3800 физиков из 182 институтов 42 стран. Современная наука принимает поистине индустриальные масштабы.

Гравитационное взаимодействие в Стандартную теорию не входит, поскольку его влияние в процессах физики частиц при достигнутых энергиях пренебрежимо мало. К тому же, квантовая теория гравитации пока не создана; о гравитационном взаимодействии мы поговорим в следующей главе.

## Контрольные вопросы

1. Перечислите известные фундаментальные взаимодействия.
2. Чем различаются два больших класса элементарных частиц — адроны и лептоны?
3. Какие частицы переносят слабые взаимодействия? В чем их отличие от фотонов?
4. На какие типы частиц подразделяются адроны?
5. Из чего состоят мезоны и барионы согласно Стандартной теории?
6. Какие частицы считаются истинно элементарными (фундаментальными)?
7. Перечислите известные ароматы кварков.
8. Что такое поколение кварков и лептонов? Сколько поколений известно?
9. Что такое цвет кварков?
10. Что такое цветовое взаимодействие?
11. Какие частицы переносят цветовое взаимодействие?
12. Что такое конфайнмент кварков?
13. Укажите экспериментальные свидетельства в пользу теории кварков.

## Глава 46

# Элементарные частицы и космология

Констатируя, что все частицы, предсказанные Стандартной моделью, уже открыты, мы неизбежно задаемся вопросом: что дальше? Логика развития науки и наблюдательные данные подсказывают, что в природе должны существовать еще неизвестные частицы и взаимодействия, так что теоретическую физику ждут, выражаясь языком А. Блока, «неслыханные перемены, невиданные мятежи».

Под логикой развития науки мы подразумеваем стремление теоретиков уменьшить число параметров Стандартной модели, значения которых определяются сейчас в экспериментах. Ясно, что провести такую программу изнутри Стандартной модели невозможно. Поэтому идет поиск более общей теории, которая объединила бы все ингредиенты Стандартной модели и в рамках единой схемы описала бы все четыре фундаментальных взаимодействия (включая гравитацию), которые в идеале предстали бы разными ипостасями единого взаимодействия.

Теории за рамками Стандартной модели предсказывают новые массивные частицы. В этой главе мы поговорим о *темной материи и темной энергии*, существование которых во Вселенной обнаруживается по их гравитационным полям, но чей состав и свойства пока являются загадкой. Возможно, решить ее помогут новые частицы. Но устоявшейся теории за пределами Стандартной модели пока нет. Предсказываемые теоретиками частицы могли бы рождаться на коллайдерах в столкновениях протонов, разогнанных до высоких энергий и потому сближающихся на очень малые расстояния. В природе подобные энергии и расстояния между частицами существовали когда-то в первые мгновения после рождения Вселенной. Таким образом, физика высоких энергий сближается с *космологией*, изучающей строение и эволюцию Вселенной как целого.

Говоря о наблюдательных данных, мы имеем в виду не только эксперименты на ускорителях, но и исследования дальнего космоса во всех доступных диапазонах электромагнитного излучения с помощью космических и наземных телескопов. Большой интерес представляет нейтринная астрономия — изучение потоков



нейтрино, прилетающих к нам из космоса или рождающихся в атмосфере. Долгожданное открытие в 2016 г. гравитационных волн создает новые перспективы получения сигналов из глубин Вселенной и ее отдаленного прошлого.

## 46.1 За пределами Стандартной модели

Стандартная модель — замечательное достижение физиков второй половины XX в. Но ее нельзя назвать совершенной теорией, так как она содержит слишком много (около двух десятков) свободных параметров, числовые значения которых определяются эмпирически, в экспериментах. Например, совершенно непонятна иерархия масс фундаментальных частиц, кварков и лептонов. Почему заряженные лептоны второго и третьего поколений, во всем похожие на электрон, намного тяжелее его: мюон — в 200 раз,  $\tau$ -лептон — в 3500 раз? Почему масса  $t$ -кварка в 90 тыс. раз больше массы  $u$ -кварка? И имеют ли массу нейтральные лептоны — нейтрино разных сортов?

### Масса нейтрино

Раньше было принято считать, что масса нейтрино равна нулю, хотя теоретических оснований к этому нет. Более того, имеются определенные экспериментальные указания на то, что масса нейтрино отлична от нуля. Если это так, то возможны так называемые *осцилляции нейтрино* (аналогичные осцилляциям нейтральных  $K$ -мезонов, описанным в гл. 45 на с. 559) — процессы взаимопревращения нейтрино разных ароматов (например,  $\nu_\mu \leftrightarrow \nu_e$ ) или нейтрино в антинейтрино ( $\nu_e \leftrightarrow \bar{\nu}_e$ ). На возможность осцилляций нейтрино указал итальянский и советский физик Б. М. Понтекорво (1957), предложивший хлор-аргонный метод регистрации этих частиц, очень слабо взаимодействующих с веществом. Взаимодействие электронного нейтрино с ядрами хлора, входящими в состав огромной массы жидкости в большом резервуаре, приводит к рождению радиоактивных ядер аргона ( $\nu_e + {}_{17}^{37}\text{Cl} \rightarrow {}_{18}^{37}\text{Ar} + e^-$ ), которые можно извлечь и «пересчитать по головам». Наблюдения солнечных нейтрино проводились американским химиком Р. Дэвисом в глубокой золотоносной шахте Хоумстейк в шт. Северная Дакота (США) в 1965–1994 гг. Из детектора, заполненного жидким тетрахлорэтиленом массой 615 тонн, Дэвис ежемесячно извлекал 20 атомов радиоактивного аргона. Количество солнечных нейтрино оказалось в 3,2 раза меньше ожидавшегося на основе общепринятой модели термоядерных реакций, происходящих в недрах нашего светила. Сейчас считается, что дефицит солнечных нейтрино был вызван их осцилляциями на пути от Солнца к Земле — превращением электронных нейтрино в мюонные, которые не регистрировались обычными детекторами. В 2002 г. Р. Дэвис и японский ученый М. Косиба (он первым в 1997 г. «поймал» несколько нейтрино, излученных при взрыве сверхновой в Большом Магеллановом облаке) были удостоены Нобелевской премии «за пионерский вклад в астрофизику, в частности за обнаружение космических нейтрино».

Нейтринная астрономия — одно из важных направлений исследования процессов, происходящих в космосе. В России работает Баксанская нейтринная обсерватория Института ядерных исследований РАН, расположенная в подземных туннелях под горой Андырчи в Приэльбрусье вблизи г. Тырныауз (Кабардино-Балкария). Для регистрации нейтрино применяется галлий-германиевый метод (предложен в 1963 г. В. А. Кузьминым), основанный на реакции  $\nu_e + {}_{31}^{71}\text{Ga} \rightarrow {}_{32}^{71}\text{Ge} + e^-$ . Галлиевая мишень состоит из 60 тонн жидкого металлического галлия высокой чистоты. Преимущество этого метода состоит в том, что он, в отличие от хлор-аргонного метода, чувствителен не только к нейтрино высоких энергий, но и к низкоэнергетическим нейтрино, испускаемым в начальной реакции  ${}^1_1p + {}^1_1p \rightarrow {}^2_1\text{H} + {}^0_1e + \nu_e$  из цепочки синтеза гелия из четырех протонов (см. разд. 44.7), являющейся основным источником энергии Солнца. Полученные экспериментальные результаты подтвердили термоядерную природу источника энергии Солнца. Они впервые показали, что дефицит солнечных нейтрино наблюдается во всем диапазоне энергий нейтрино. Это закрыло попытки решения проблемы дефицита нейтрино путем модификации модели Солнца и привело к однозначному выводу о существовании нейтринных осцилляций.

В настоящее время осцилляции нейтрино наблюдались не только в нейтринном потоке, исходящем от Солнца, но и у так называемых атмосферных нейтрино, рожденных в столкновении космических лучей с атомами воздуха, а также в экспериментах на реакторах и ускорителях. В 2015 г. канадец А. Макдональд и японец Т. Кадзита получили Нобелевскую премию «за открытие нейтринных осцилляций, показывающее, что нейтрино имеют массу».

Стандартная теория без особого труда позволяет включить массы нейтрино в используемую сейчас схему. Эксперименты указывают, что масса нейтрино не превышает  $0,01 - 0,1 \text{ эВ}/c^2$ .

## Варионная асимметрия Вселенной

О несохранении  $CP$ -четности в распадах нейтральных  $K$ - и  $B$ -мезонов мы уже говорили. Стандартная теория позволяет формально описать нарушение  $CP$ -симметрии, но природа явления не совсем ясна. Очевидно, что нарушение  $CP$ -симметрии носит фундаментальный характер и непосредственно связано с какими-то глобальными закономерностями. Поэтому открытие этого явления потрясло физиков, тем более что оно до сих пор остается одной из загадок природы. В частности, из нарушения  $CP$ -симметрии следует также нарушение  $T$ -симметрии (от англ. *time* — время), т. е. симметрии при отражении времени. Такая симметрия означает, что если возможно какое-то движение системы, то возможно и обратное движение, когда система проходит те же состояния, но в обратном порядке. На макроскопическом уровне стрела времени необратима, но закон возрастания энтропии имеет статистический характер, а на уровне простых физических систем  $T$ -симметрия не вызывала сомнений. Но одним из фундаментальных свойств любой квантовой теории поля является строго доказанная ***СРТ-теорема***, утверждающая симметрию относительно всех трех

операций — зарядового сопряжения, зеркального отражения пространства и отражения времени. Поэтому нарушение  $CP$ -инвариантности практически однозначно свидетельствует о нарушении и  $T$ -симметрии, т. е. о выделенности одного из направлений течения времени.

Существует еще одна загадка природы, которая, возможно, связана с нарушением  $CP$ -симметрии, — так называемая **барионная асимметрия Вселенной**. Наблюдения показывают, что в видимой части Вселенной отсутствуют сколько-нибудь крупные образования из антивещества. Если бы антивещество присутствовало в значительных количествах, то непременно дало бы о себе знать излучением, возникающим при аннигиляции с обычным веществом. В то же время античастицы во всем, кроме своих зарядов, подобны частицам. При образовании Вселенной в Большом взрыве число античастиц должно было бы быть равно числу частиц. Почему же сейчас во Вселенной вещество явно превалирует над антивеществом?

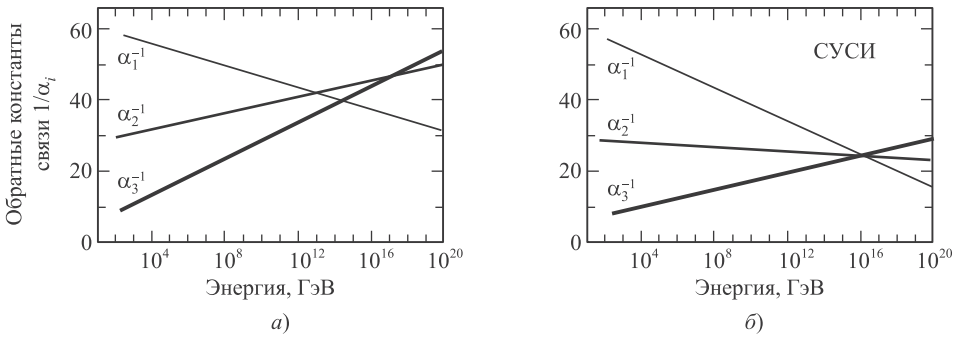
В момент Большого взрыва температура была настолько велика, что массы ультрарелятивистских частиц особой роли не играли и все они мало чем отличались от фотонов. По мере расширения и остывания Вселенной более тяжелые барионы «отсоединились» от фотонов и стали интенсивно аннигилировать друг с другом, порождая новые фотоны. В настоящее время на каждый барион приходится более миллиарда фотонов реликтового излучения, оставшихся от той эпохи. И почти не сохранилось антибарионов — они все были уничтожены в процессах аннигиляции. С ними вместе исчезло такое же число барионов, а наша Вселенная (и мы с вами) построена из остатка. Это означает, что изначально число барионов на одну миллиардную часть превышало число антибарионов. Но как возник этот избыток? Согласно гипотезе А. Д. Сахарова (1967), ответственность лежит на явлении нарушения  $CP$ -симметрии, которое приводит к зарядовой асимметрии продуктов ядерных реакций.

Но из барионной асимметрии Вселенной следует вывод о нарушении закона сохранения барионного числа. Если сначала число барионов было равно числу антибарионов, то полный барионный заряд Вселенной равнялся нулю. А в нашу эпоху он явно отличен от нуля, а барионное число сохраняется в процессах распада и рождения частиц. В частности, протон стабилен, потому что ему не на что распасться: более легкого бариона не существует.

Общепринятых объяснений наблюдаемой барионной асимметрии Вселенной нет. Попытки построить так называемые теории **Великого объединения** естественным образом ведут к нарушению сохранения барионного заряда и к нестабильности протона. Но существует и экзотическая гипотеза, что темная материя состоит из антибарионов, которые не взаимодействуют с обычным веществом. Тогда полный барионный заряд Вселенной может быть равен нулю.

## Великое объединение

Теориями **Великого объединения** называли теоретические схемы на основе калибровочных полей Янга—Миллса, которые в рамках единых общих принципов



**Рис. 46.1.** Поведение бегущих констант связи в Стандартной модели (а) и в суперсимметричной теории (б): точки пересечения прямых указывают на область энергий, когда соответствующие взаимодействия объединяются в одно целое

описали бы электрослабые и цветовые взаимодействия. Трём фундаментальным силам соответствуют три группы симметрий, о которых мы говорили выше: цветовая неабелева симметрия  $SU(3)$ , приводящая к сильному взаимодействию, неабелева симметрия  $SU(2)$  между фермионами того же поколения, приводящая к их слабому взаимодействию, и абелева симметрия  $U(1)$ , порождающая электромагнитное взаимодействие. Интенсивность этих трех фундаментальных взаимодействий определяется соответствующими безразмерными **константами связи**  $\alpha_i$ , которые меняются с энергией (**бегущие константы связи**). На рисунке 46.1, а показана расчетная зависимость от энергии обратных значений этих трех констант: электромагнитной  $\alpha_1 = 5\alpha_{em}/(3\cos^2\theta_W)$ , слабой  $\alpha_2 = \alpha_{em}/\sin^2\theta_W$  и сильной  $\alpha_3^1$ . Точки пересечения прямых показывают области энергий, где интенсивности соответствующих взаимодействий становятся одинаковыми. Говоря иначе, при таких энергиях уже невозможно отличить эти взаимодействия друг от друга (а кварки от лептонов). Электромагнитное взаимодействие полностью объединяется со слабым при  $E \sim 10^{13}$  ГэВ, а слабое объединяется с сильным при  $E \sim 10^{17}$  ГэВ. Но если соединить теории Великого объединения с идеями **суперсимметрии (СУСИ)**, то все три прямые пересекутся в одной точке при энергии  $E \sim 10^{16}$  ГэВ (рис. 46.1, б). Эти расчеты явились сильным стимулом к развитию **суперсимметричных теорий**.

## Суперсимметрия

В стандартной модели резко различаются бозоны — кванты поля с целым спином (фотон  $\gamma$ ,  $W^+$ ,  $W^-$  и  $Z^0$ , глюоны  $g$ ) и фермионы — частицы вещества с полуцелым спином (кварки и лептоны). Внутри фермионов существует определенная симметрия, проявляющаяся в группировке частиц по парам, но с бозонами эта

<sup>1</sup>Напомним, что  $\alpha_{em} = e^2/(4\pi\epsilon_0\hbar c)$  — постоянная тонкой структуры КЭД, равная  $1/137$  при малых энергиях, а  $\theta_W$  — угол Вайнберга в теории электрослабых взаимодействий, см. (45.5).

симметрия никак не связана. Очень интересная теория, называемая *суперсимметрией*, объединяет в пары частицы разных спинов — 0, 1/2, 1, 3/2, 2. Подход дает принципиальную возможность построить квантовую теорию гравитации и объединить электрослабые и сильные взаимодействия<sup>2</sup>.

Стандартная модель не обладает суперсимметрией, но может быть расширена до так называемой *минимальной суперсимметричной Стандартной модели (МССМ)* путем включения новых полей. Теория предсказывает, что каждой фундаментальной частице Стандартной модели должен соответствовать суперпартнер, спин которого меньше на 1/2. Гипотетические частицы со спином 0, являющиеся *суперпартнерами* обычных кварков и лептонов, называют *скварками* и *слептонами* соответственно. Предсказывается существование и суперпартнеров переносчиков взаимодействий — *фотино*, *вино*, *зино* и *глюино* со спином 1/2. Бозон Хиггса также имеет суперпартнера — *хиггсино* со спином 1/2.

Легчайшие из суперпартнеров могли бы быть составляющими темной материи. Пока нет никаких экспериментальных данных, подтверждающих существование суперсимметрии в природе. Однако суперсимметрия привлекает внимание теоретиков, прилагающих усилия по построению единой теории фундаментальных взаимодействий.

## Суперструны

Теоретики повели также атаку на еще более амбициозную цель — создание теории *Всего сущего* (от англ. *Theory of Everything*), где все четыре фундаментальных взаимодействия (включая гравитацию) слились бы в рамках единой квантовой теории поля. Если в теориях суперсимметричного Великого объединения бегущие константы пересекаются в одной точке при энергиях  $E \sim 10^{16}$  ГэВ, то гравитация все еще остается в этой области слишком слабой и может присоединиться к тем трем взаимодействиям лишь при энергиях  $E \sim 10^{19}$  ГэВ. Таким образом, мы уже представляем, где квантовая гравитация должна влиться в единую теорию всех фундаментальных взаимодействий и дело за малым — создать такую объединенную теорию.

Венцом усилий теоретиков в этом направлении можно считать *теорию суперструн* в 10-мерном пространстве. В этой теории основными структурными единицами являются не точечные частицы, как в Стандартной модели, но протяженные объекты — струны и мембраны. Из десяти измерений шесть свернуты в тонкие трубки радиусом порядка планковской длины (см. на с. 588). Поэтому мы не замечаем этих измерений, довольствуясь оставшимися нам четырьмя — временем и трехмерным пространством.

Математически возможны пять теорий суперструн — слишком много для единой теории поля. Но есть сильные подозрения, что все они — это лишь

---

<sup>2</sup>Более подробно на доступном уровне о суперсимметрии можно прочитать в научно-популярной книге: Г. Кейн. Суперсимметрия и не только. От бозона Хиггса к новой физике. М.: «Бином. Лаборатория знаний», 2015.

эквивалентные версии единой  $M$ -теории в 11-мерном пространстве, где свернуто уже семь измерений. Математика, лежащая в основе такой теории, настолько сложна, что один из ее авторов и активных пропагандистов, гуру современной физики Э. Уиттен сказал, что «это наука 21-го века, случайно открытая в 20-м». Как бы то ни было, выход за пределы Стандартной теории и создание Теории всего сущего — дело ученых текущего столетия. Задача трудная, но цель великая: понять, почему наше пространство-время такое, какое оно есть, узнать, почему и как произошел Большой взрыв (в теории суперструн существует минимальный размер, меньше которого Вселенная стать не может, — возможно, так решится проблема начальной сингулярности), обнаружить закономерности в иерархии масс фундаментальных частиц и многое другое. Об этих сложных теориях увлекательно написано в научно-популярной книге<sup>3</sup>.

## 46.2 Вселенная — прошлое и будущее

Даже самые нелюбопытные люди слышали об уникальном явлении — рождении нашей Вселенной в процессе, который мы называем *Большим взрывом* и детали которого нам пока неизвестны. Свет от далеких объектов приходит к нам с большим запаздыванием. Поэтому, изучая дальний космос, астрофизики заглядывают в прошлое Вселенной. Но в прошлом Вселенная была очень горячей и очень плотной. Поэтому, изучая ее прошлое, ученые пытаются добыть сведения о свойствах микромира на таких малых расстояниях, которые недоступны современным ускорителям.

### Расширяющаяся Вселенная и закон Хаббла

Многие столетия звездное небо представлялось нашим предкам символом неизменности и покоя. Земля вращается вокруг Солнца, одни созвездия на небе сменяются другими, но через год все приходит на круги своя, и так из века в век. Прецессия земной орбиты меняет вид звездного неба, но ее период составляет 25 765 лет, так что она (прецессия) отчетливо проявляется лишь в масштабах тысячелетий.

Говоря научным языком, Вселенная в целом представлялась ученым стационарной. Более того, они давно поняли, что у Вселенной не может быть центра (или, что то же самое, любая ее точка может быть взята за центр). В 1584 г. Джордано Бруно, поклонник и последователь коперниковой системы мира, сказал: «Вселенная едина, бесконечна, неподвижна. <...> Мы наверняка можем утверждать, что вся Вселенная есть целиком центр, или что центр Вселенной

---

<sup>3</sup>Б. Грин. Элегантная Вселенная. Суперструны, скрытые размерности и поиски окончательной теории. М.: Едиториал УРСС, 2005.

повсюду, и что окружность не имеется ни в какой части»<sup>4</sup>. Выражаясь современным языком, Вселенная представлялась однородной и изотропной — ее свойства в среднем должны быть одинаковы во всех точках и по всем направлениям.

Гелиоцентрическая система и закон всемирного тяготения навели полный порядок в Солнечной системе: планеты ходили по своим орбитам, как по часам, а когда выявлялись отклонения, то астрономы обнаруживали источник возмущения (так, на кончике пера, были открыты Нептун и Плутон). Но в применении к Вселенной в целом ньютоновская механика приводила к неразрешимым трудностям. Рядом ученых было замечено, что небо ночью не должно быть черным. Если Вселенная бесконечна, то куда бы мы ни посмотрели, в любом небольшом телесном угле окажутся мириады и мириады звезд. Число звезд, находящихся в этом телесном угле, растет как куб расстояния, а поток излучения от них падает как квадрат расстояния. Поэтому, казалось бы, светимость неба при бесконечной однородной Вселенной должна стремиться к бесконечности. На самом деле это не так, даже если не учитывать поглощение света межгалактическими газовыми и пылевыми облаками. Звезды просто загораживают друг друга, и поэтому даже при бесконечной Вселенной свет на Землю приходит от ограниченного числа звезд, находящихся на конечном расстоянии от нас.

Для оценки этого расстояния вспомним наши рассуждения о ежике в тумане и атомной бомбе (см. т. 1, разд. 14.1). Там была выведена формула (14.14) для оценки видимости в тумане, которую вполне можно применить к оценке радиуса видимости в космосе:

$$R \approx \frac{12}{n\pi d^2}, \quad (46.1)$$

где  $d$  — средний диаметр звезды, а  $n$  — концентрация звезд в мировом пространстве. В качестве  $d$  возьмем диаметр вполне заурядной звезды — нашего Солнца:  $d = 1,4 \cdot 10^9$  м. Концентрация звезд в окрестностях Солнца оценивается как  $n \approx 3 \cdot 10^{-50}$  м<sup>-3</sup>. Отсюда получаем оценку радиуса видимости в бесконечной Вселенной:  $R \approx 10^{32}$  м  $\approx 10^{16}$  св. лет. Эта величина на много порядков превышает размеры видимой части Вселенной (13,7 млрд св. лет), но она не бесконечна. Отсюда можно сделать вывод, что блеск любой точки на небе должен быть таким же, как и блеск средней звезды. Этот вывод известен как **фотометрический парадокс**, или **парадокс Ольберса** — по имени обсуждавшего его в 1823 г. немецкого астронома Г. В. Ольберса (хотя тот и не был первым, кто обратил внимание на проблему, но это уже парадоксы науки). Парадокс Ольберса противоречит нашему каждодневному опыту, но тем не менее этот парадокс оставался неразрешенным вплоть до XX в., пока не обнаружили, что Вселенная расширяется и ее возраст конечен. Из самого вывода вышеупомянутой формулы

<sup>4</sup> Дюк. Бруно. Пять диалогов о причине, начале и едином. М.: Госполитиздат, 1949. С. 273. Любопытно, что к подобной формуле в конце XII в. пришли богословы: «Бог есть умопостигаемая сфера, центр коей находится везде, а окружность нигде». В литературе средневековья эта мысль неоднократно повторялась. И даже в середине XVII в. Б. Паскаль в своих «Мыслях» привел ту же формулу: «Природа — это бесконечная сфера, центр которой везде, а окружность нигде». (Утверждается, правда, что в рукописи вместо слова «бесконечная» сначала было написано, а потом зачеркнуто слово «устрашающая».)

(14.14) следует, что отношение видимого размера Вселенной к вычисленному нами значению  $R$  дает долю площади небесной сферы, занятой звездами, и она оказывается ничтожно малой:  $\delta S/S \approx 10^{-22}$ .

Еще большие трудности возникают при попытке применения математического аппарата классической механики к описанию Вселенной в целом. Вспомним уравнение (15.35) в т. 1, связывающее напряженность электростатического поля с потенциалом:

$$\vec{E} = -\vec{\nabla}\varphi. \quad (46.2)$$

Подставим отсюда выражение для  $\vec{E}$  в первое уравнение Максвелла (см. (23.33) в т. 1):

$$\operatorname{div}\vec{E} = \vec{\nabla} \cdot \vec{E} = \frac{\rho_e}{\varepsilon_0}. \quad (46.3)$$

В результате получаем уравнение Пуассона для электростатического потенциала

$$\Delta\varphi = -\frac{1}{\varepsilon_0}\rho_e. \quad (46.4)$$

Символ  $\Delta$  означает здесь оператор Лапласа:

$$\Delta = \vec{\nabla}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (46.5)$$

Для точечного заряда  $q$ , расположенного в начале координат, плотность заряда имеет вид  $\rho_e = q\delta(\vec{r})$ , а решением уравнения Пуассона

$$\Delta\varphi = -\frac{1}{\varepsilon_0}q\delta(\vec{r}) \quad (46.6)$$

будет хорошо известный потенциал Кулона  $\varphi = q/(4\pi\varepsilon_0 r)$ .

Закон Кулона во всем подобен закону всемирного тяготения Ньютона и переходит в него при замене электрического заряда  $q$  на массу  $m$ , а выражения  $1/(4\pi\varepsilon_0)$  — на гравитационную постоянную  $-G$  (отрицательный знак здесь нужен для смены знака потенциала: одноименные электрические заряды отталкиваются, а массы притягиваются). Как нетрудно видеть, при такой замене плотность электрического заряда  $\rho_e$  переходит в обычную плотность  $\rho$ , а уравнение Пуассона для потенциала гравитационного поля принимает вид

$$\Delta\varphi = 4\pi G\rho. \quad (46.7)$$

Для гравитационного поля точечной массы  $M$ , когда  $\rho = M\delta(\vec{r})$ , отсюда следует гравитационный потенциал  $\varphi = -GM/r$  (ср. т. 1, разд. 6.5). Для существования решений уравнения (46.7) плотность вещества на больших расстояниях должна падать быстрее, чем  $1/r^2$ . В противном случае следующее из (46.7) выражение для гравитационного потенциала

$$\varphi = -G \int \frac{\rho}{r} dV$$

приводит к расходящимся интегралам. Этот вывод известен как **гравитационный парадокс Зеелигера** по имени обсуждавшего его еще в 1895 г. немецкого



астронома Х. фон Зеелигера. В этом смысле ньютоновская Вселенная конечна и имеет выделенную точку — центр сферического распределения вещества, вдали от которого плотность быстро падает до нуля. Если же применить ньютоновскую механику к бесконечной однородной и изотропной Вселенной, то градиент потенциала надо положить равным нулю во *всех точках* (иначе в ней на пробное тело будет действовать сила в каком-то выделенном направлении). Тогда будет равна нулю и вторая производная от потенциала (т. е. левая часть уравнения (46.7), а с ней вместе — и правая часть этого уравнения). Мы приходим к действительно парадоксальному выводу, что статическая однородная изотропная Вселенная должна быть абсолютно пустой ( $\rho = 0$ ). Поистине, гипотеза о бесконечной статической однородной Вселенной оказалась устрашающей. По счастью, окружающий нас мир не пуст, да и небо ночью не сияет как звезды. Стало быть, тем хуже для теории, хотя и хорошо проверенной в масштабах нашей планеты и Солнечной системы.

В 1915–1916 гг. Эйнштейн сформулировал уравнения Общей теории относительности (ОТО), в которой гравитация ассоциировалась с кривизной пространства-времени. Траектории тела (планеты, астероида или кометы) в поле тяготения Солнца стали интерпретироваться как свободное движение в пространстве-времени, искривленном из-за присутствия огромной массы нашей звезды. Теория Эйнштейна предсказала красное гравитационное смещение — уменьшение частоты излученного звездой фотона вследствие потери энергии при преодолении им гравитационного притяжения звезды (или, что то же самое, замедления хода времени в гравитационном поле). Этот эффект был подтвержден в лабораторных экспериментах в 1960 г. (см. т. 1, разд. 9.13). Предсказанное ОТО отклонение луча света в гравитационном поле было впервые измерено в 1919 г. во время солнечного затмения. Одним из достижений ОТО также стало объяснение вращения перигелия Меркурия (примерно  $43''$  в столетие), обнаруженного еще в середине XIX в. и остававшегося долгое время единственным астрономическим фактом, противоречащим классической механике. В применении к этому случаю ОТО позволила вычислить поправку к ньютоновскому закону всемирного тяготения, что и объяснило наблюдаемое явление.

Рассматривая трудности применения ньютоновской теории тяготения к Вселенной, Эйнштейн начал с модификации уравнения Пуассона (46.7), предложив записать его как

$$\Delta\varphi - \Lambda\varphi = 4\pi G\rho, \quad (46.8)$$

где  $\Lambda$  представляла бы собой некоторую универсальную постоянную размерностью  $[L^{-2}]$ . Тогда при постоянной плотности  $\rho_0$  существовало бы решение уравнения (46.8)  $\varphi_0 = -4\pi G\rho_0/\Lambda$ . Такой мир, как отмечал Эйнштейн, «не имел бы центра по отношению к гравитационному полю и не было бы надобности допускать, что плотность уменьшается на бесконечности; наоборот, и средний потенциал и средняя плотность были бы постоянны вплоть до бесконечности»<sup>5</sup>.

<sup>5</sup> А. Эйнштейн. Собрание научных трудов. М.: Наука, 1965. Т. 1, с. 604.

Для точечной массы  $M$  в начале координат уравнение (46.8) ведет к так называемому *потенциалу Юкавы*

$$\varphi(r) = -\frac{GM}{r}e^{-\sqrt{\Lambda}r}. \quad (46.9)$$

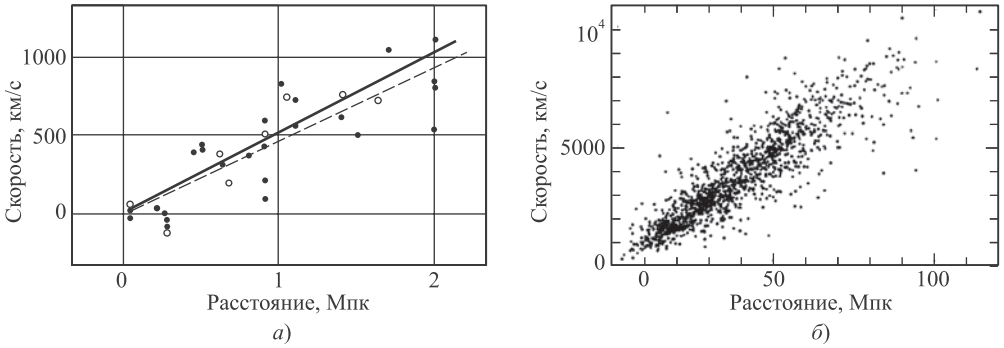
Подобную модификацию закона всемирного тяготения Ньютона предлагал еще Зеелигер и другие астрономы. При малых значениях  $\Lambda \ll 1/R_0^2$ , где  $R_0$  — радиус Солнечной системы, отклонение потенциала (46.9) от стандартного ньютоновского  $-MG/r$  не приведет к наблюдаемым следствиям в движении планет, комет и астероидов.

Понимая, что ОТО в определенном приближении переходит в ньютоновскую теорию тяготения и, следовательно, столкнется с теми же проблемами при описании стационарной и однородной Вселенной, в 1917 г. Эйнштейн поступил с уравнениями ОТО точно так же, как и при переходе от (46.7) к (46.8): он просто «руками» ввел в свои уравнения дополнительное слагаемое — *космологический член*, пропорциональный *космологической постоянной*  $\Lambda$ . Тензорные дифференциальные уравнения ОТО гораздо сложнее уравнения Пуассона, но идея их модификации была абсолютно такой же. На этом пути Эйнштейн нашел решения модифицированных уравнений ОТО, описывающих пространственно замкнутый мир конечного объема (но не имеющий границ), заполненный равномерно распределенным веществом.

Вселенная Эйнштейна — это четырехмерный цилиндр: ось времени — прямая, время не искривляется, а пространство сферически симметрично, причем его кривизна не зависит от времени. В том же 1917 г. голландский физик В. де Ситтер нашел другое решение уравнений с космологическим членом, соответствующее четырехмерной сфере: пространственная часть такая же, как и у мира Эйнштейна, но «искривленное» время течет иначе. Мир де Ситтера свободен от вещества, его стационарность поддерживается только лямбда-членом (энергией вакуума). В обоих этих стационарных Вселенных параметр  $\Lambda$  обратно пропорционален квадрату радиуса кривизны пространства-времени.

В 1922 г. российский и советский математик и физик А. А. Фридман нашел нестационарные решения изначальных уравнений ОТО без космологического члена. Эти решения описывали либо расширяющуюся, либо сжимающуюся Вселенную, которая при этом была однородной (плотность вещества в ней в любой момент времени была одинаковой во всех точках пространства). Фридман показал, что миры Эйнштейна и де Ситтера — это единственные стационарные решения уравнений ОТО с космологическим членом и что они являются частными случаями найденных им решений. Работа, присланная Фридманом из Петрограда, была опубликована в самом важном физическом журнале того времени — немецком *Zeitschrift für Physik*, она привлекла внимание и вызвала возражения Эйнштейна, который в 1923 г. нашел ошибку в своих вычислениях и в итоге признал правоту Фридмана.

Однако теория теорией, но снова все дело решили данные наблюдений. За десять лет до нестационарных решений Фридмана американец В. Слайфер, изучавший спектр излучения галактик, обнаружил красное смещение, которое было



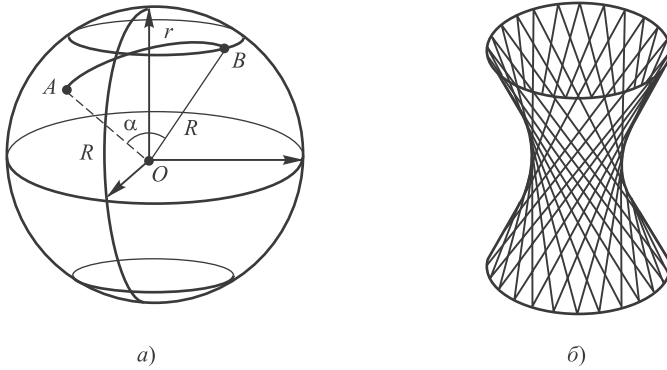
**Рис. 46.2.** Зависимость скорости удаления галактик от расстояния: *а)* график из работы Хаббла (1929); *б)* современные (2010) данные по 1355 галактикам (отклонения точек от прямой линии связаны с погрешностью измерений и локальным собственным движением галактик); расстояния даны в мегапарсеках (1пк = 3,26 св. лет)

интерпретировано как эффект Доплера вследствие удаления галактик от Земли. Но в 1929 г. американский астроном Э. Хаббл опубликовал ключевую работу с описанием зависимости радиальных скоростей галактик на основе собственных наблюдений и данных других астрономов, где обнаружил удивительное явление. Во-первых, всюду наблюдалось красное смещение и нигде — синее. Иными словами, все галактики, как сговорившись, удалялись от Земли, и никто из них не приближался. Во-вторых, красное смещение спектральных линий галактик было тем больше, чем дальше от Земли они находилась. Иначе говоря, чем дальше находилась галактика, тем выше была скорость ее удаления от Земли. Эти результаты не зависели от направления, под которым была видна галактика. Обработка наблюдательных данных привела Хаббла к эмпирическому закону, который теперь называется **законом Хаббла**:

$$\vec{u} = H\vec{r}. \quad (46.10)$$

Здесь  $r$  — расстояние до галактики,  $u$  — ее радиальная скорость, а  $H$  — коэффициент пропорциональности с размерностью обратного времени ( $[H]=[1/T]$ ), который называют теперь **постоянной Хаббла**. Результаты Хаббла представлены на рис. 46.2, *а*. Для численной оценки постоянной Хаббла сам автор получил значение, завышенное в полтора раза. Современное (2015) значение  $H = 67,74 \pm 0,46$  (км/с)/Мпк (см. рис. 46.2, *б*).

Рассмотрим в качестве наглядного примера замкнутый сферический двумерный мир, т.е. обычную сферу (рис. 46.3). Представим себе, что на этой сфере живут обитатели, которые не воспринимают третьего измерения, хотя тот факт, что они живут на кривой поверхности, не мог ускользнуть от их внимания. Для них кратчайшее расстояние  $r$  между точками  $A$  и  $B$  (аналог прямой в плоском пространстве) — это дуга большого круга, т.е. линия пересечения сферы с плоскостью, проведенной через точки  $A$ ,  $B$  и центр кривизны их мира  $O$ . Это расстояние можно выразить через радиус кривизны  $R = OA = OB$  и угол  $\alpha$  между радиусами, проведенными из центра  $O$  в точки  $A$  и  $B$ :  $r = \alpha R$ . Представим теперь, что радиус кривизны зависит от времени



**Рис. 46.3.** а) Сфера — модель мира с постоянной положительной кривизной: кратчайшее расстояние  $r$  между точками  $A$  и  $B$  на сфере радиусом  $R$  ( $R = OA = OB$ ) — это дуга большого круга, т. е. линия пересечения сферы с плоскостью, проведенной через точки  $A$ ,  $B$  и центр сферы  $O$ ; б) гиперboloид — модель мира с постоянной отрицательной кривизной

(например, увеличивается). Тогда будет увеличиваться и расстояние  $r$  между точками  $A$  и  $B$ . Дифференцируя уравнение связи  $r$  и  $R$  по времени, получаем аналогичное соотношение  $\dot{r} = \alpha \dot{R}$ . Теперь мы можем связать расстояние  $r$  и скорость его изменения  $u = \dot{r}$ , исключив вспомогательный параметр — угол  $\alpha$ :

$$u = Hr, \quad H = \frac{\dot{R}}{R}. \quad (46.11)$$

Параметр  $H$  (постоянная Хаббла в нашей простой модели) зависит, вообще говоря, от времени, но одинаков для любых двух точек на сфере. Ясно, что выделенной точки на самой сфере нет. Наблюдатель в точке  $A$  увидит, как увеличивается размер его мира, как от него удаляется точка  $B$  и все остальные окружающие объекты, причем чем они дальше, тем больше скорость удаления. Но точно такую же картину увидит и наблюдатель в точке  $B$ . Житель такой двумерной Вселенной (местный Паскаль) мог бы сказать, что его мир — это круг, центр которого находится везде, а окружность (граница) — нигде.

Следствием закона Хаббла был вывод, что убегают от Земли не галактики, а расширяется само пространство, что увеличивается объем наблюдаемой Вселенной. Этот поразительный результат был в согласии с теоретическими предсказаниями Фридмана (который до этого открытия не дожил). Закон Хаббла означает, что в прошлом пространство было меньше. Если учесть, что  $1 \text{ Мпк} = 3,1 \cdot 10^{19} \text{ км}$ , то постоянная Хаббла  $H = 2,195 \cdot 10^{-18} \text{ с}$ . Обратная величина  $1/H = 0,456 \cdot 10^{18} \text{ с} = 14,5 \text{ млрд лет}$  с точностью до численных множителей порядка единицы дает оценку времени, прошедшего с начала расширения Вселенной. Современное значение возраста Вселенной близко к этой оценке и равно 13,7 млрд лет.

Теория Фридмана, подкреплённая законом расширения Хаббла, решила парадокс Ольберса. Мы не живем в бесконечной и статической Вселенной.

У Вселенной есть конечный возраст, и она расширяется. После начала расширения прошло всего лишь 13,7 млрд лет, и мы можем видеть только звезды, находящиеся от нас на конечных расстояниях 13,7 св. лет. Это означает, что для наблюдения доступно конечное число звезд. Из-за того что скорость света конечна, существуют области Вселенной, которые мы не видим, потому что у света просто не было времени долететь до Земли<sup>6</sup>. Но с каждым мгновением доступная наблюдению область Вселенной увеличивается.

Космологическое красное смещение происходит из-за того, что, пока световая волна путешествует от галактики к Земле, пространство растягивается и вместе с ним увеличивается длина световой волны, т. е. уменьшается ее частота. Чем длиннее путь, тем больше времени затрачивается на него и тем сильнее за время путешествия растягиваются пространство и длина волны.

### Критическая плотность

Рассмотрим расширение Вселенной, заполненной веществом — космической пылью и газом. Мы не станем, как следовало бы, использовать сложные уравнения ОТО, а применим классический закон всемирного тяготения Ньютона. На этом этапе результаты, получаемые в рамках этих теорий, совпадают. Пусть внутри сферы радиусом  $R$  содержится сферически симметричное распределение вещества с плотностью  $\rho$  и общей массой  $M$ . Поскольку нас интересует расширение самого пространства, а не локальное движение небесных тел, будем считать, что при изменении радиуса выделенного нами объема масса внутри него не меняется. Ускорение частицы, находящейся на поверхности сферы, определяется вторым законом Ньютона и законом всемирного тяготения:

$$\frac{d^2R}{dt^2} = -G \frac{M}{R^2} = -\frac{4\pi}{3} G \rho R. \quad (46.12)$$

Во втором равенстве мы выразили массу  $M$  через плотность ее распределения  $\rho$ . Используя закон расширения Хаббла  $dR/dt = HR$ , вторую производную в (46.12) можно представить в виде

$$\frac{d(HR)}{dt} = \frac{dH}{dt}R + \frac{dR}{dt}H = \left( \frac{dH}{dt} + H^2 \right) R. \quad (46.13)$$

Мы видим, что и в левой, и в правой частях уравнения (46.12) стоит выражение, пропорциональное  $R$ . Таким образом, мы пришли к выражению, куда уже не

<sup>6</sup>Американский поэт-романтик Эдгар По в 1848 г. интуитивно угадал возможное объяснение парадокса Ольберса: наблюдаемая нами Вселенная конечна, потому свет мог не успеть дойти до нас из самых отдаленных ее точек. В своей поэме в прозе «Эврика» По писал (перевод К. Бальмонта): «Если бы непрерывность звезд была бесконечна, тогда бы заднее поле неба являло нам единообразную светящуюся, подобную исходящей от Млечного пути, — ибо безусловно не было бы точки, на всем этом заднем поле, где не существовало бы звезды. Единственный способ поэтому, при таком положении вещей, понять пустоты, что открывают наши телескопы в бесчисленных направлениях, предположить, что рассеяние от незримого заднего поля так несметно, что ни один его луч доселе совершенно не мог нас достигнуть. Что это может быть так, кто решится отрицать?»

входят ни радиус  $R$ , ни масса  $M$ :

$$\boxed{\frac{dH}{dt} = -H^2 - \frac{4\pi}{3}G\rho.} \quad (46.14)$$

Обратимся снова к уравнению (46.12) и умножим обе его части на  $dR/dt$ . Тогда (46.12) можно записать в виде равенства полных производных:

$$\frac{1}{2} \frac{d}{dt} \left( \frac{dR}{dt} \right)^2 = GM \frac{d}{dt} \left( \frac{1}{R} \right). \quad (46.15)$$

Отсюда следует закон сохранения энергии:

$$\frac{1}{2} \left( \frac{dR}{dt} \right)^2 - \frac{GM}{R} = E. \quad (46.16)$$

Полную энергию  $E$  (константу интегрирования) можно выразить через значения  $H_0$ ,  $\rho_0$ ,  $R_0$  (учтем, что  $M = 4\pi\rho_0/R_0^3$ ) в произвольный момент времени  $t_0$  (например, в современную эпоху). Тогда  $E$  можно записать в виде

$$E = \left( \frac{1}{2}H_0^2 - \frac{4\pi}{3}G\rho_0 \right) R_0^2. \quad (46.17)$$

Из уравнения (46.16) следует, что при отрицательном значении параметра  $E < 0$  скорость расширения  $\dot{R}$  обратится в нуль при достижении Вселенной максимального радиуса  $R_{\max} = GM/|E|$ , после чего расширение сменится сжатием. Этот случай соответствует миру конечного объема с положительной постоянной кривизной. Если же  $E \geq 0$ , то Вселенная будет расширяться вечно. При  $E > 0$  это будет мир с отрицательной кривизной, а при  $E = 0$  — плоский евклидов мир с нулевой кривизной. Таким образом, критическое значение плотности  $\rho_{\text{кр}}$ , при котором происходит смена этих режимов, определяется из условия  $E = 0$ . Из уравнения (46.17) следует тогда, что

$$\rho_{\text{кр}} = \frac{3H_0^2}{8\pi G} = 8,6 \cdot 10^{-27} \text{ кг/м}^3. \quad (46.18)$$

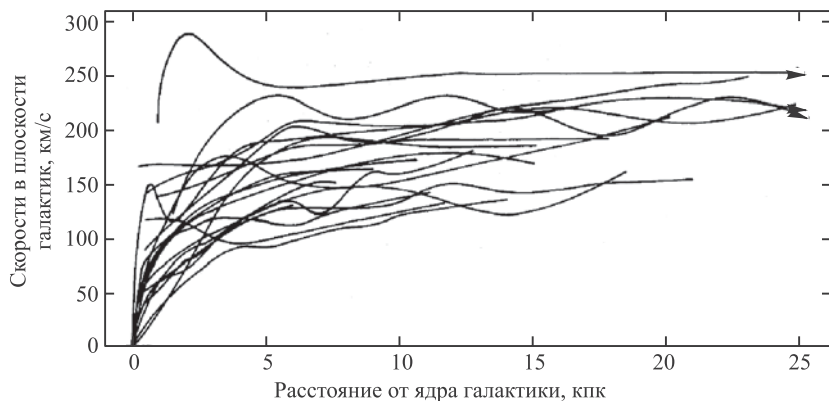
Плотность видимого («светящегося») вещества во Вселенной оценивается примерно в  $\rho_{\text{вещ}} = 10^{-28} \text{ кг/м}^3$ . Таким образом, отношение  $\rho_{\text{вещ}}/\rho_{\text{кр}} \approx 0,01$ , т. е. видимое вещество обеспечивает всего лишь около 1% критической плотности. Это должно бы свидетельствовать, что геометрия нашего мира неевклидова и он имеет отрицательную кривизну. Но геометрия Вселенной, судя по всему, все-таки является евклидовой. Учет других видов обычного вещества поднимает его долю до 5%, но его плотность все равно остается гораздо меньше критической плотности.

## Темная материя

Согласно законам Кеплера, чем дальше находится планета от Солнца, тем больше период ее вращения, т. е. тем меньше ее скорость. Год на Марсе длится 687 земных дней, а на Юпитере — 12 земных лет. Приступая к изучению

движения галактик, астрономы предполагали, что скорости вращения входящих в них звезд и других объектов будут подчиняться той же закономерности, что наблюдается у планет. Однако и здесь природа преподнесла сюрприз. В 1933 г. американский астроном Ф. Цвикки измерял относительные скорости галактик, входящих в Скопление Волос Вероники, — группу примерно из тысячи скоординированно движущихся галактик, находящихся от Земли на расстоянии около 320 млн св. лет. Цвикки обнаружил, что видимого (светящегося) вещества не хватает для объяснения скорости относительного движения галактик. Он выдвинул гипотезу о существовании во Вселенной *скрытой массы* (теперь ее называют *темной материей*), проявляющей себя лишь по гравитационному взаимодействию. По оценке Цвикки, скрытая масса в 400 раз превышала массу видимых объектов.

Около 1959 г. астрономы разработали спектроскопические методы измерения скорости вращения рукавов спиральных галактик (таких, как Туманность Андромеды или наша собственная галактика Млечный путь). В середине 1970 г. американский астроном Вера Рубин детально изучила вращение многих спиралевидных галактик. Было обнаружено, что скорости звезд и пылевых облаков, расположенных вдали от центра галактик, не уменьшаются с расстоянием, а остаются почти постоянными. Рубин доказала, что наблюдаемые особенности могут быть объяснены наличием скрытой массы, существование которой предположил еще Цвикки. На рисунке 46.4 из работы Рубин показаны измеренные ею кривые вращения для двадцати одной галактики. Сейчас принято считать, что темная материя сосредоточена в массивном *галактическом гало*, имеющем сферическую форму и окружающем любую галактику. Светимость гало практически равна нулю, лишь изредка попадающиеся там облака газа позволяют определить скорость их обращения на различных расстояниях от центра галактики.



**Рис. 46.4.** Кривые вращения двадцати одной спиральной галактики (из работы В. Рубин, 1980); расстояния показаны в килопарсеках

Чтобы понять эти результаты, рассмотрим простую модель сферической галактики. Обозначим через  $M^*(R)$  суммарную массу обычного вещества и темной материи, распределенных сферически симметричным образом внутри сферы радиусом  $R$ . Скорость вращения объекта массой  $m$  в гравитационном поле такого распределения масс получим из закона всемирного тяготения Ньютона и второго закона Ньютона:

$$\frac{mv^2}{R} = G \frac{mM^*(R)}{R^2}, \quad (46.19)$$

откуда

$$v = \sqrt{\frac{GM^*(R)}{R}}. \quad (46.20)$$

Если темной материи нет, а вся масса  $M$  обычного вещества сосредоточена в силовом центре (как в Солнечной системе), то мы приходим к третьему закону Кеплера (см. разд. 6.1) для вращения планет. Скорость обращения планеты в зависимости от ее расстояния до Солнца дается в этом случае формулой

$$v = \sqrt{\frac{GM}{R}}. \quad (46.21)$$

На рисунке 46.5 показана кривая зависимости (46.21) скорости обращения планет от расстояния до Солнца, а также лежащие на нее точки, соответствующие наблюдаемым данным о скорости вращения планет Солнечной системы.

Пусть теперь обычное вещество распределено в галактике сферически симметричным образом с плотностью  $\rho(r)$ . В качестве простейшего примера возьмем распределение по линейному закону:

$$\rho(r) = \begin{cases} \rho_0 \left(1 - \frac{r}{R_0}\right), & r \leq R_0, \\ 0, & r \geq R_0. \end{cases} \quad (46.22)$$

Здесь  $R_0$  — радиус галактики, а  $\rho_0$  — плотность вещества в ее центре. Массу  $M$  галактики найдем, интегрируя плотность  $\rho(r)$  по всему пространству. В сферических координатах интеграл по углам дает  $4\pi$ , и мы получаем выражение

$$M = 4\pi \int_0^{R_0} \rho(r)r^2 dr = 4\pi\rho_0 \int_0^{R_0} \left(1 - \frac{r}{R_0}\right)r^2 dr = \frac{\pi}{3}\rho_0 R_0^3. \quad (46.23)$$

Выражая из (46.23) плотность  $\rho_0$  через массу  $M$ , приходим к формуле

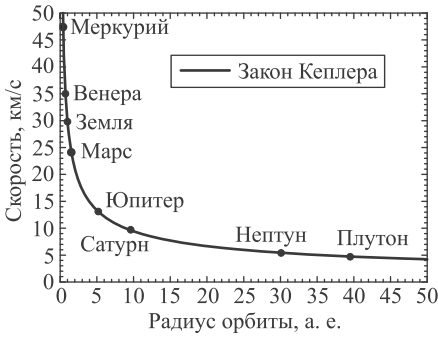
$$\rho(r) = \frac{3M}{\pi R_0^3} \left(1 - \frac{r}{R_0}\right). \quad (46.24)$$

Массу  $M(R)$  внутри сферы радиусом  $R < R_0$  найдем, интегрируя плотность (46.24) по объему этой сферы:

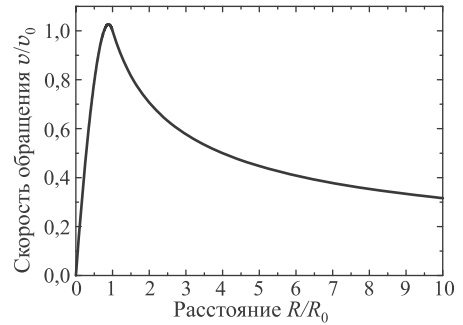
$$M(R) = 4\pi \int_0^R \rho(r)r^2 dr = \frac{12M}{R_0^3} \int_0^R \left(1 - \frac{r}{R_0}\right)r^2 dr = 4M \left(\frac{R}{R_0}\right)^3 \left(1 - \frac{3}{4} \frac{R}{R_0}\right). \quad (46.25)$$

На рисунке 46.6 показана зависимость скорости обращения от расстояния, рассчитанная по формуле (46.20), где мы для  $M^*(R)$  при  $R \leq R_0$  взяли выражение (46.25) и обозначили  $v_0 = \sqrt{GM/R_0}$ . При  $R \geq R_0$  имеем  $M(R) = M$ , а для скорости обращения —





**Рис. 46.5.** Зависимость скорости обращения планет от их расстояния до Солнца



**Рис. 46.6.** Скорость обращения объекта при радиусе галактики  $R_0$  и линейной плотности распределения вещества в ней

выражение (46.21). При удалении от центра увеличение скорости обусловлено ростом гравитирующей массы внутри сферы радиусом  $R$ . При больших расстояниях рост массы прекращается и скорость начинает падать по тому же закону, по которому падает скорость планет.

Постоянство наблюдаемой скорости вращения галактик свидетельствует, что с ростом расстояния  $R$  объекта от центра галактики притягивающая его масса  $M^*(R)$  увеличивается по линейному закону, как это следует из (46.20). Такой вклад и дает темная материя в галактическом гало.

Если темная материя распределена в гало с плотностью  $\rho_{\text{ТМ}}(r)$ , то масса темной материи внутри сферы радиусом  $R$  равна

$$M_{\text{ТМ}}(R) = 4\pi \int_0^R \rho_{\text{ТМ}}(r)r^2 dr. \quad (46.26)$$

Дифференцируя по  $R$ , можно выразить плотность темной материи через ее массу внутри сферы радиусом  $R$ :

$$\rho_{\text{ТМ}}(R) = \frac{M'_{\text{ТМ}}(R)}{4\pi R^2}. \quad (46.27)$$

Если обозначить символом  $v_\infty$  постоянную скорость обращения объектов за пределами видимой части галактики, то масса обычного вещества и темной материи внутри сферы радиусом  $R$  будет равна

$$M^*(R)|_{R \gg R_0} = \frac{v_\infty^2}{G} R. \quad (46.28)$$

Так как за пределами галактики масса  $M$  обычного вещества в ней уже не меняется, а масса  $M_{\text{ТМ}}(R)$  линейно растет, мы будем считать, что (46.28) дает оценку массы темной материи. Тогда, подставляя (46.28) в (46.27), находим поведение плотности темной материи на больших расстояниях:

$$\rho_{\text{ТМ}}(R)|_{R \gg R_0} = \frac{v_\infty^2}{4\pi G R^2}. \quad (46.29)$$

С учетом (46.29) оценим плотность темной материи вблизи внешнего края галактик, используя данные Рубин с рисунка 46.4. Возьмем для примера крайнюю правую точку на верхней кривой, для которой  $v_\infty = 250$  км/с =  $2,5 \cdot 10^5$  м/с

при  $R = 25$  кпк. Учитывая, что  $1 \text{ пк} = 3,08 \cdot 10^{16} \text{ м}$ , имеем  $R = 7,7 \cdot 10^{20} \text{ м}$ . Значение ньютоновской гравитационной постоянной  $G = 6,67 \cdot 10^{-11} \text{ м}^3 \text{ кг}^{-1} \text{ с}^{-2}$ . Подставляем численные данные в (46.29) и получаем оценку

$$\rho_{\text{ТМ}} = \frac{(2,5 \cdot 10^5)^2}{4\pi \times 6,67 \cdot 10^{-11} \times (7,7 \cdot 10^{20})^2} \approx 10^{-22} \frac{\text{кг}}{\text{м}^3}. \quad (46.30)$$

Много это или мало? Сравним наш результат со средней плотностью Солнечной системы. Вся ее масса практически сосредоточена в Солнце:  $M_{\odot} = 2 \cdot 10^{30} \text{ кг}$ . За радиус  $R_0$  Солнечной системы примем *гелиопаузу* — расстояние  $R_0 = 90 \text{ а. е.} = 1,3 \cdot 10^{13} \text{ м}$ , на котором гравитационное влияние Солнца уже не превосходит влияния звезд. При таком выборе границы Солнечной системы ее объем составляет  $V_0 = 4\pi R_0^3/3 \approx 10^{40} \text{ м}^3$ , а ее плотность  $\rho_S = M_{\odot}/V \approx 2 \cdot 10^{-10} \text{ кг/м}^3$ . Эта величина на много порядков больше плотности темной материи. Средняя плотность обычного вещества в типичных галактиках составляет  $\rho_G \approx 2 \cdot 10^{-21} \text{ кг/м}^3$ , т. е. она уже одного порядка с плотностью темной материи. А плотность вещества в скоплении галактик  $\rho_{GG} \approx 7 \cdot 10^{-25} \text{ кг/м}^3$  уже гораздо меньше плотности темной материи. Отсюда следует, что в масштабах Солнечной системы темная материя никакой роли не играет, она проявляется на галактических и межгалактических расстояниях.

Результаты Цвикки, Рубин и других астрономов убедили физиков в существовании темной материи. Пока это всего лишь красивое название для того, чего мы не знаем: темная материя обнаруживает себя через гравитационное взаимодействие и иначе никак не дает о себе знать. Вопрос «всего лишь» в том, чтобы понять, из чего она состоит. Для изучения Вселенной современные астрономы используют электромагнитное излучение самых разных диапазонов: радиоволны, инфракрасное, рентгеновское и гамма-излучения, но скрытая масса в галактиках и скоплениях галактик остается невидимой. Каков бы ни был ее состав, темная материя не участвует ни в сильном, ни в электромагнитном взаимодействии (в противном случае ее бы обнаружили не только по гравитационному воздействию). По современным оценкам на долю темной материи приходится около 23% массы Вселенной.

## Темная энергия

Введенный Эйнштейном лямбда-член можно учесть, заменив ньютоновское уравнение движения (46.12) на уравнение

$$\frac{d^2 R}{dt^2} - \frac{\Lambda c^2}{3} R = -\frac{4\pi}{3} G \rho R. \quad (46.31)$$

Нетрудно видеть, что лямбда-член эквивалентен введению силы, расталкивающей Вселенную и препятствующей ее гравитационному сжатию. Если перенести в правую часть уравнения (46.31) лямбда-член, то его можно интерпретировать как добавку к плотности обычного вещества:

$$\frac{1}{R} \frac{d^2 R}{dt^2} = -\frac{4\pi}{3} G \left( \rho - \frac{\Lambda c^2}{4\pi G} \right). \quad (46.32)$$

Эта добавка отрицательна, как будто мы имеем дело с какими-то отрицательными массами, т. е. с антигравитацией. Но не будем спешить.

Присутствие лямбда-члена в правой части уравнения (46.32) современные ученые связывают с влиянием пустого пространства (вакуума). После знакомства с квантовой механикой не должно вызвать удивления утверждение, что вакуум далеко не пуст. Скажем, даже при отсутствии электромагнитных полей в вакууме могут рождаться виртуальные фотоны. Если разместить параллельно другу другу две незаряженные металлические пластины, то из-за граничных условий (равенство нулю электрического поля на их поверхности) электромагнитное поле между ними будет представлять собой стоячую волну, т. е. содержать лишь дискретный набор частот. Из-за отсутствия части спектра плотность энергии электромагнитного поля между пластинами будет меньше, чем плотность энергии снаружи. Соответственно давление вакуумных флуктуаций снаружи пластин будет превосходить давление изнутри, и на пластины будет действовать сближающая их сила (*эффект Казимира*). Сила Казимира заметна лишь на небольших расстояниях и зависит от формы проводников — для других конфигураций она может сменить знак и превратиться в силу отталкивания.

Гравитационное поле тоже должно порождать вакуумные флуктуации, но их свойства совершенно необычны. Обозначим символом  $w_\Lambda$  плотность энергии вакуума и будем считать, что эта величина постоянна во всем пространстве Вселенной. Такое предположение выглядит естественно, так как гравитация универсальна и всепроникающа. Но это значит, что при увеличении какого-то объема его вакуумная энергия увеличивается на величину  $dE_\Lambda = w_\Lambda dV$ . По закону сохранения энергии такое увеличение должно быть равно работе сил давления  $pdV$ . Отсюда следует, что *положительная* плотность энергии вакуума создает *отрицательное* давление  $p_\Lambda = -w_\Lambda$ . Субстанцию с такими необычными свойствами называли *темной энергией*.

В уравнении (46.12) давление нигде не фигурировало, поскольку для обычного вещества оно ничтожно мало. Но для излучения, например, это уже не так, и из ОТО следует, что мир Фридмана описывается уравнением

$$\frac{d^2 R}{dt^2} = -\frac{4\pi}{3}G \left( \rho + \frac{3p}{c^2} \right) R. \quad (46.33)$$

Присутствие темной энергии мы должны учесть, подставив вместо  $p$  отрицательное давление вакуума  $p_\Lambda$  и приняв во внимание наличие массовой плотности темной энергии  $\rho_\Lambda = w_\Lambda/c^2$ . В результате получаем уравнение движения

$$\frac{d^2 R}{dt^2} = -\frac{4\pi}{3}G \left( \rho + \rho_\Lambda - \frac{3p_\Lambda}{c^2} \right) R = -\frac{4\pi}{3}G \left( \rho - \frac{2w_\Lambda}{c^2} \right) R. \quad (46.34)$$

Сравнивая уравнения (46.32) и (46.34), находим связь плотности энергии вакуума с космологической постоянной:

$$\Lambda = \frac{8\pi G}{c^4} w_\Lambda. \quad (46.35)$$

Если в какой-то период эволюции Вселенной в ней доминирует темная энергия, то одним из решений уравнения (46.34), где мы положим  $\rho = 0$ , будет экспоненциальное расширение Вселенной:

$$R \propto \exp\left(\sqrt{\frac{\Lambda c^2}{3}} t\right). \quad (46.36)$$

Дифференцируя (46.36) и применяя закон Хаббла (46.10), получаем оценку для космологической постоянной:

$$\Lambda = 3 \left(\frac{H}{c}\right)^2 \approx 1,6 \cdot 10^{-52} \text{ м}^{-2}. \quad (46.37)$$

В рамках современных представлений о расширении Вселенной наблюдательные данные приводят к близкому значению  $\Lambda = 1,19 \cdot 10^{-52} \text{ м}^{-2}$ .

Величина  $R_\Lambda = 1/\sqrt{\Lambda} \approx 10^{26} \text{ м} \approx 10^{10} \text{ св. лет}$  близка к размеру наблюдаемой части Вселенной. Иными словами, наличие космологической постоянной не скажется на движении небесных тел не только в масштабах Солнечной системы, но даже в масштабах галактик и их скоплений.

Уравнение (46.35) позволяет теперь оценить плотность энергии вакуума:

$$\begin{aligned} w_\Lambda &= \frac{\Lambda c^4}{8\pi G} = 5,8 \cdot 10^{-10} \text{ Дж/м}^3 = 3,6 \text{ ГэВ/м}^3, \\ \rho_\Lambda &= w_\Lambda/c^2 = \frac{\Lambda c^2}{8\pi G} \approx 6,4 \cdot 10^{-27} \text{ кг/м}^3. \end{aligned} \quad (46.38)$$

Используя (46.18), получаем для отношения плотности темной энергии к критической плотности:  $\rho_\Lambda/\rho_{\text{кр}} \approx 0,74$ . И действительно, по современным наблюдательным данным доля темной энергии в составе Вселенной составляет 72%.

Заметим, что стационарность миров Эйнштейна и де Ситтера обеспечивалась при определенных соотношениях между космологической постоянной  $\Lambda$  и постоянным радиусом кривизны системы. При другом значении  $\Lambda$  эти миры также будут нестационарными.

Следует подчеркнуть, что введение в теорию лямбда-члена не означает появления антигравитации: силы тяготения по-прежнему притягивают тела друг к другу, но в присутствии среды (вакуума) с необычными свойствами — отрицательным давлением — массовая плотность всех видов энергии эффективно уменьшается и даже меняет знак, что и приводит к смене знака гравитационного потенциала: из потенциала притяжения он превращается в потенциал отталкивания. Все зависит от уравнения состояния среды — зависимости давления  $p$  от плотности энергии  $w = \rho c^2$ , которое можно записать в виде  $p = \eta w$ . Например, для электромагнитного излучения  $\eta = 1/3$  и обычное положительное давление не ослабляет гравитацию, а усиливает ее — тоже ведь парадоксальный вывод. Для среды с отрицательным давлением при  $-1/3 < \eta < 0$  гравитация ослабляется, как следует из уравнения (46.33), но экспоненциального расширения Вселенной

не возникает. При  $\eta = -1/3$  мы приходим к статическому решению. Такой зависимости давления от плотности энергии соответствует значение космологической постоянной

$$\Lambda = \frac{4\pi G\rho}{c^2}, \quad (46.39)$$

что совпадает с условием статичности мира Эйнштейна (см. (46.32)). Наконец, при отрицательном давлении среды и  $\eta < -1/3$  мы получаем экспоненциальное расширение Вселенной. Для давления вакуума, как мы видели,  $\eta = -1$ . Именно это давление, а вовсе не антигравитация «растягивает» нашу Вселенную.

В 1998 г. было обнаружено, что галактики, находящиеся к нам ближе и потому видимые в более «современную» эпоху Вселенной, разбегаются быстрее, чем очень отдаленные галактики, которые мы наблюдаем в гораздо более ранний период истории Вселенной. Этот вывод подтвержден исследованиями реликтового излучения и крупномасштабной структуры Вселенной. В 2011 г. астрофизики С. Перлмуттер, Б. Шмидт и А. Рисс получили Нобелевскую премию «за открытие ускоренного расширения Вселенной посредством наблюдения дальних сверхновых». Считается, что за ускорение расширения Вселенной ответственна **темная энергия**, хотя ее существование нельзя обнаружить непосредственно традиционными методами наблюдений: как и в случае темной материи, о присутствии темной энергии можно судить косвенно по гравитационному воздействию на обычное вещество.

## Гравитация и планковские масштабы

Гравитация была первым фундаментальным взаимодействием, с которым столкнулся человек. Открытие Ньютоном закона всемирного тяготения позволило понять, что падение тел на Землю и движение планет в Солнечной системе обусловлены одними и теми же силами.

Гравитационное взаимодействие управляет наиболее глобальными процессами во Вселенной, в частности, обеспечивая стабильность нашей Солнечной системы. Оно много слабее других видов взаимодействий в случае частиц малой массы, для которых существенны квантовые эффекты. Поэтому для таких частиц не обнаружены пока никакие экспериментальные проявления гравитационного взаимодействия. Слабость гравитационного взаимодействия в микромире становится очевидной из сравнения характеризующей его безразмерной константы, построенной из ньютоновской гравитационной постоянной  $G$  и других фундаментальных постоянных. В законе всемирного тяготения Ньютона сила притяжения между двумя телами (в качестве таковых мы возьмем основную структурную единицу атомного ядра — протоны) выражается как  $F = Gm_p^2/r^2$ . Произведение силы на квадрат расстояния  $F_g r^2$  имеет размерность Дж · м; точно такую же размерность имеет комбинация фундаментальных постоянных  $\hbar c$ . Поэтому величина  $\alpha_g = Gm_p^2/\hbar c$  безразмерна и может служить для характеристики интенсивности гравитационного взаимодействия. Численное значение этой **константы гравитационного взаимодействия** равно  $\alpha_g = 5.9 \cdot 10^{-39}$ . Сравнивая с аналогичными параметрами  $\alpha_{em} \approx 1/137$  в КЭД и  $\alpha_w \approx 10^{-5}$  в

теории слабых взаимодействий, убеждаемся, что гравитационное взаимодействие на 33 порядка менее интенсивно, чем слабое взаимодействие, и на 36 порядков слабее электромагнитного. Поэтому последнее доминирует в микросистемах с размерами порядка атомов и молекул, да и в процессах сильного и слабого взаимодействий, где характерные расстояния гораздо меньше, гравитация совершенно несущественна. Однако гравитация, несомненно, играла важнейшую роль в первые мгновения после начала расширения Вселенной, которая тогда имела исчезающе малые размеры.

Из фундаментальных постоянных  $\hbar$  и  $c$  нельзя построить величину размерности массы. Но это позволяет сделать третья фундаментальная постоянная —  $G$ , входящая в закон всемирного тяготения Ньютона:

$$M_{Pl} = \sqrt{\frac{\hbar c}{G}}. \quad (46.40)$$

Эта величина называется **планковской массой**. В масштабах микромира планковская масса огромна. Подставляя в (46.40) численные значения фундаментальных констант, находим:  $M_{Pl} = 1,22 \cdot 10^{19}$  ГэВ/ $c^2 = 2,18 \cdot 10^{-8}$  кг.

Мы построили планковскую массу формально, исходя из соображений размерности. Но можно усмотреть в ней и физический смысл. Если попытаться локализовать частицу массой  $m$  в области размером порядка гравитационного радиуса  $\Delta x \sim R_g = Gm/c^2$  и учесть, что неопределенность ее импульса в этом ультрарелятивистском случае равна по порядку величины  $\Delta p \sim mc$ , то из соотношения неопределенностей  $\Delta x \Delta p \sim \hbar$  следует нижняя оценка на массу такой частицы:  $m \geq M_{Pl}$ . Таким образом, планковская масса — верхний предел для массы частицы (при больших массах она коллапсирует в *черную дыру*) и нижний предел для массы черной дыры.

Огромны и следующие отсюда **планковская энергия** и **планковская температура**:

$$\begin{aligned} E_{Pl} &= M_{Pl}c^2 = \sqrt{\frac{\hbar c^5}{G}} = 1,2 \cdot 10^{19} \text{ ГэВ}, \\ T_{Pl} &= E_{Pl}/k_B = 1,4 \cdot 10^{32} \text{ К}. \end{aligned} \quad (46.41)$$

Такие колоссальные энергии и температуры достигались лишь в первые мгновения после Большого взрыва.

С помощью планковской массы нетрудно получить также **планковскую длину** (как комптоновскую длину волны объекта с планковской массой), **планковское время** (за которое свет преодолевает планковскую длину), **планковскую плотность энергии** и **планковскую массовую плотность**:

$$\begin{aligned} l_{Pl} &= \hbar/(M_{Pl}c) = \sqrt{\hbar G/c^3} = 1,6 \cdot 10^{-35} \text{ м}, \\ t_{Pl} &= l_{Pl}/c = \sqrt{\hbar G/c^5} = 5,4 \cdot 10^{-44} \text{ с}, \\ w_{Pl} &= E_{Pl}/l_{Pl}^3 = 0,29 \cdot 10^{124} \text{ ГэВ/м}^3, \\ \rho_{Pl} &= M_{Pl}/l_{Pl}^3 = 0,52 \cdot 10^{97} \text{ кг/м}^3. \end{aligned} \quad (46.42)$$

Эти величины указывают на пространственно-временные размеры, в которых классическая теория (ОТО) уже заведомо не будет работать и должна быть

заменена на квантовую гравитацию. Ниже указанных пределов мы не можем говорить о пространстве и времени в привычных для нас терминах. Поэтому пока невозможно сказать, что происходило в первые  $10^{-44}$  с после Большого взрыва, и сам вопрос представляется некорректным. Существование сингулярности в «нулевой момент» есть следствие применения классической теории (ОТО), но на планковских временах существенную роль должны были играть квантовые флуктуации гравитационного поля (по сути, самого пространства-времени). Теория квантовой гравитации, как уже говорилось, еще не создана. В приближении слабого поля Эйнштейн предсказал существование гравитационных волн, которые и были обнаружены в 2016 г. (см. разд. 29.7). Квантование этих волн по обычным правилам приводит к заключению, что переносчиком гравитационных взаимодействий должна быть безмассовая частица со спином  $s = 2$  — *гравитон*.

Казалось бы, планковские масштабы должны регулировать и свойства темной энергии, коль скоро последняя связана с гравитационными флуктуациями вакуума. Но «естественное» планковское значение для величины размерности обратного квадрата длины равно  $\Lambda_{Pl} = 1/l_{Pl}^2 \approx 0,38 \cdot 10^{70} \text{ м}^{-2}$ , в то время как космологическая постоянная  $\Lambda = 1,6 \cdot 10^{-52} \text{ м}^{-2} \approx 10^{-122} \Lambda_{Pl}$ . Пока неясно, каким образом физика частиц может объяснить появление столь малого множителя (122 порядка!) в космологии.

## Горячая Вселенная и Большой взрыв

Расширение Вселенной от очень малых (планковских) размеров до нынешних еще не означало, строго говоря, Большого взрыва. Одной из проблем астрофизики первой половины XX в. было объяснение распространенности химических элементов во Вселенной. В 1930-х гг. уже было известно, что во Вселенной преобладает водород. Химические элементы могут рождаться в ядерных реакциях *первичного нуклеосинтеза* из протонов и нейтронов, но если Вселенная живет в стационарном состоянии неограниченно долго, то в конечном итоге в ней должны остаться лишь самые устойчивые элементы из середины таблицы Менделеева — железо в первую очередь, а это не так, как следовало из данных наблюдений.

Впервые идея связи первичного нуклеосинтеза с теорией Фридмана была высказана в статье «Расширяющаяся Вселенная и происхождение элементов» (1946) эмигрировавшего в США советского физика Г. А. Гамова, бывшего студента Фридмана. В 1948 г. Гамов и его аспирант Р. Альфер (также из семьи эмигрантов из России) предложили теорию «горячей Вселенной», в которой учитывалось, что расширяющаяся Вселенная не находится в термодинамическом равновесии и имеет конечное время жизни. Начальным состоянием они предложили считать не только очень плотное, но и очень горячее вещество. В одной из научно-популярных телепередач этот процесс был назван *Большим взрывом*, и это название прижилось и у публики, и у физиков. В том же году Альфер в своей диссертации дал более детальную картину нуклеосинтеза и пришел к выводу, что первая фаза в образовании химических элементов —

синтез дейтерия — заняла примерно 200–300 секунд. Современная наука считает, что вся последующая судьба Вселенной определилась в первые три минуты. Теория горячей Вселенной и Большого взрыва развивалась и уточнялась, и сейчас общепринят взгляд, что собственно в Большом взрыве родились водород, гелий и несколько легких элементов (литий, бериллий, бор). Вся средняя часть таблицы Менделеева от углерода до примерно железа образовалась в ядерных реакциях, происходящих в недрах звезд, а более тяжелые элементы от железа до урана и плутония появились при взрывах сверхновых.

Дальнейшими достижениями теория горячей Вселенной обязана работам сотрудников Лаборатории прикладной физики университета Д. Хопкинса (г. Балтимор, США). В том же 1948 г. Альфер и его коллега Р. Герман предсказали существование реликтового излучения. При сверхвысоких температурах в молодой Вселенной ультрарелятивистские частицы мало чем отличались от фотонов, и электромагнитное излучение находилось в равновесии с веществом, а его спектр совпадал со спектром абсолютно черного тела (АЧТ). По мере расширения и остывания Вселенной энергия частиц уменьшалась, и когда температура упала примерно до 3000 К, стало возможным образование нейтральных атомов водорода и гелия (эпоха рекомбинации), и фотонный газ «отсоединился» от вещества. Это произошло примерно через 300–400 лет после начала расширения Вселенной. В последующие миллиарды лет электромагнитному излучению, сохранившемуся от эпохи рекомбинации, оставалось лишь медленно остывать, сохраняя форму спектра АЧТ. Альфер и Герман пришли к выводу, что это *реликтовое излучение* (или *космический микроволновой фон*) должно сохраниться до наших дней и иметь температуру около 5 К. Этот вывод был подтвержден наблюдениями радиоизлучения из космоса в 1965 г. (по современным измерениям температура реликтового излучения составляет  $2,72548 \pm 0,00057$  К). В работе 1953 г. Альфер, Герман и Фоллин рассчитали распространенность химических элементов во Вселенной и получили результаты, близкие к современным: около 75% водорода и 25% гелия. Прочие элементы, из которых построены планеты и мы с вами, составляют очень незначительную долю состава Вселенной.

Большой взрыв, случившийся 13,7 млрд лет тому назад, — это не обычный взрыв, который мы способны представить в нашем воображении при таких словах. При химическом взрыве осколки разлетаются в разные стороны из одного центра, при Большом взрыве центра Вселенной не было, как нет его и сейчас. Просто она очень быстро увеличивалась в размерах. Основные этапы эволюции Вселенной, соответствующие современным представлениям, показаны на рис. 46.7.

По прошествии эпохи квантовых флуктуаций пространства-времени гравитация отделилась от прочих фундаментальных взаимодействий — можно сказать, что через  $10^{-44}$  с после Большого взрыва родились наше пространство и время.

Далее пошла фаза объединения кварк-глюонной плазмы в протоны и нейтроны. К моменту времени порядка  $10^{-35}$  с, когда Вселенная была размером меньше протона, сильные взаимодействия отделились от электрослабых. И тут произошло событие, которое можно считать невероятным: расширение Вселенной



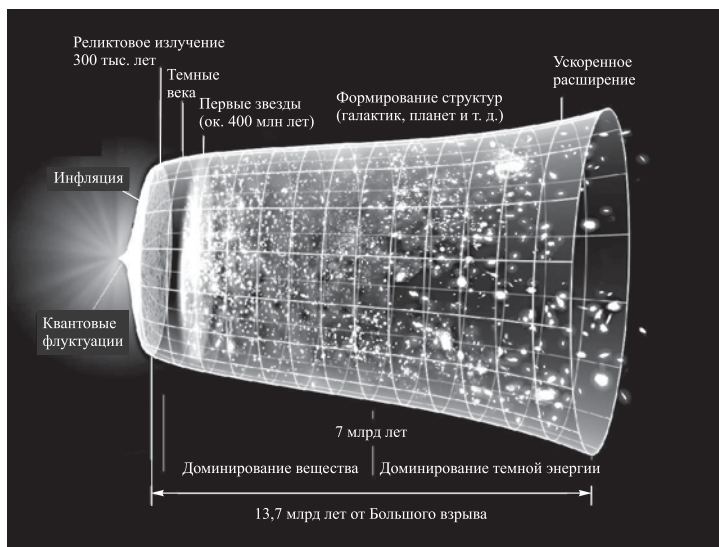


Рис. 46.7. Основные этапы эволюции Вселенной

резко ускорилось. За время около  $10^{-32}$  с ее размер увеличился на 50–100 порядков, его уже можно было сравнить с размером яблока. Эта фаза называется *инфляцией Вселенной* (от англ. *inflation* — раздувание, надувание). Подобное экспоненциальное расширение мы уже видели в мире де Ситтера. Различные механизмы инфляции были предложены в 1980-х гг. разными учеными (среди основоположников чаще всего упоминаются имена советских и экс-советских физиков А. Старобинского, А. Линде, В. Муханова и американца А. Гута). Теория инфляции объяснила ряд фактов, остававшихся непонятными в прежнем сценарии Большого взрыва (в частности, тот факт, что температура реликтового излучения оказалась одинаковой во всех частях Вселенной, даже находящихся на столь больших расстояниях, что они никак не могли повлиять друг на друга)<sup>7</sup>. Одни считают теорию инфляции научной фантастикой, другие предрекают ее авторам Нобелевские премии. Словом, будущее покажет.

За инфляционной фазой развития последовала фаза *радиационного доминирования*, оставившая нам в наследство реликтовое излучение.

Фаза доминирования вещества, начавшаяся с эпохи рекомбинации, продолжалась около 7 млрд лет, а потом, судя по последним открытиям, сменилась фазой доминирования темной энергии.

<sup>7</sup>В теории инфляции становится ясно, что все эти удаленные области Вселенной просто «помнят» о начальных условиях, когда они находились в непосредственном контакте в юной Вселенной размером с маковое зернышко и даже еще меньше. Инфляция объясняет также малую (или вообще нулевую) кривизну нынешнего пространства, которое близко к плоскому евклидовому пространству классической физики Ньютона — точно так же поверхность очень большой сферы кажется ее обитателям плоской.

Согласно астрофизическим наблюдениям, Вселенная имеет нулевую кривизну, т. е. мы все-таки живем в плоском мире, описываемом, грубо говоря, обычной геометрией Евклида. Однако, как мы уже знаем, такое возможно, если плотность всех видов энергии во Вселенной имеет критическое значение. Пока все имеющиеся наблюдательные данные свидетельствуют в пользу так называемой *космологической  $\Lambda$ CDM-модели* (произносится «Лямбда-Си-Ди-Эм», аббревиатура от англ. *Cold Dark Matter* — холодная темная материя). В ее рамках по современным данным, как уже говорилось, вклад темной энергии в критическую плотность Вселенной составляет около 72%, темное вещество вносит примерно 23%, а на долю обычного вещества, которое мы, собственно, и изучали в этом учебнике, остается всего 5%. Получается, что физика — самая продвинутая из всех естественных наук — едва-едва выходит из состояния младенчества: за все прошедшие века она в какой-то степени изучила всего 5% состава Вселенной.

И никто не может ответить на вопрос: какой конец ожидает нашу Вселенную? Имеются варианты, и каждый может выбрать тот, который ему больше по вкусу.

- **Большое сжатие.** Если плотность всех видов энергии превышает критическое значение, нынешнее расширение когда-нибудь остановится и сменится сжатием. Вселенная вернется в первоначальное состояние немислимой плотности и температуры, после чего (кто знает?) может последовать новый Большой взрыв и родится новая Вселенная. Пока все данные указывают на то, что плотность Вселенной не превышает критического значения. Но мы не знаем природы темной энергии, и нет никакой уверенности, не сменится ли нынешнее ускоренное расширение сжатием.
- **Большой разрыв.** Если будет продолжаться экспоненциальное расширение Вселенной с ускорением, то удаленные объекты могут исчезнуть за горизонтом событий: свет от них перестанет до нас доходить, и наступит момент, когда наша галактика останется в одиночестве в видимой нами части Вселенной и никакой связи с ушедшими за горизонт объектами уже не будет. Потом за горизонт начнут уходить звезды, пока в одиночестве не останется одна только Солнечная система. И такое сокращение видимой части Вселенной будет продолжаться, пока она не достигнет субатомных размеров. Объекты размера больше расстояния до горизонта просто будут разорваны на куски: сначала планеты, потом атомы, наконец элементарные частицы. Большой разрыв преодолет даже силы, связывающие кварки и глюоны в адронах. Дата такого сокрушительного события в истории Вселенной точно не назначена — быть может, через 22 млрд лет, считая от данного момента.
- **Увядание Вселенной.** Тихо и спокойно Вселенная продолжает расширяться, она пока еще вовсе не стара, большинство звезд только-только зажглось. Примерно через 100 трлн лет они все догорят, исчерпав свое горючее. Гравитация начнет собирать остатки вещества в черные дыры, а что будет потом, когда черные дыры испарятся, даже гадать уже невозможно.

Мы привели несколько сценариев, у космологов имеются и другие, но захочет ли природа играть в пьесах, написанных людьми?

## Контрольные вопросы

1. Что такое барионная асимметрия Вселенной?
2. Что такое суперсимметрия и какие последствия она может иметь для «зоопарка» фундаментальных частиц?
3. Укажите основные нерешенные проблемы физики элементарных частиц.
4. Сформулируйте закон Хаббла.
5. Галактики удаляются от нас во всех направлениях. Почему нельзя утверждать, что Земля находится в центре мироздания?
6. Что такое космологическое красное смещение и чем оно отличается от гравитационного красного смещения и эффекта Доплера?
7. Как обнаружили темную материю?
8. Что считается свидетельством присутствия темной энергии во Вселенной?
9. Перечислите основные этапы эволюции Вселенной.

# ЭПИЛОГ

*А для низкой жизни были числа,  
Как домашний, подъяремный скот,  
Потому что все оттенки смысла  
Умною число передает.*

Николай Гумилев. «Слово»

Начав с простейших механических систем, давно и хорошо известных людям, мы дошли в этом курсе до глубин Вселенной, о которых, наоборот, нам известно до обидного мало. И, раздумывая над тем, как закончить учебник, мы решили поиграть в игру «А что, если бы...?» А что, если бы величина гравитационной постоянной (или скорости света, или постоянной Планка, или массы электрона и протона и т. д.) была бы чуть больше (или меньше) того значения, которое мы узнали из своих экспериментов? Если говорить коротко (а для длинного и вразумительного разговора у нас уже не остается места), то общепринятое мнение таково: даже при сравнительно небольших отклонениях фундаментальных констант и других физических параметров от известных нам значений эволюция Вселенной и разных объектов в ней была бы совсем иной, причем такой, что возникновение жизни вряд ли было бы возможным. Это явление называется ***тонкой настройкой параметров Вселенной.***

Что случилось бы, например, если бы масса электрона увеличилась всего лишь в 4 раза и составила бы  $m_e = 2 \text{ МэВ}/c^2$  (при сохранении неизменными всех прочих констант)? Поскольку масса протона  $m_p = 938,3 \text{ МэВ}/c^2$ , а нейтрона  $m_n = 939,6 \text{ МэВ}$ , разность их масс равна  $m_n - m_p = 1,3 \text{ МэВ}/c^2$ . Значит, нейтрон стал бы стабильной частицей и не смог бы, как сейчас, распадаться по слабому взаимодействию на протон, электрон и антинейтрино ( $n \rightarrow p + e^- + \bar{\nu}_e$ ) вследствие закона сохранения энергии. Зато стала бы возможной реакция поглощения электрона протоном:  $p + e^- \rightarrow n + \nu_e$ . Это значит, что атомы водорода стали бы нестабильными и обычное вещество, в том числе звезды и планеты, не смогло бы существовать. Соответственно во Вселенной, наполненной лишь нейтронами, нейтрино и фотонами, жизнь в известной нам форме невозможна.

Энергия связи ядра гелия ( $\alpha$ -частицы) составляет около 7 МэВ, а энергия электростатического отталкивания протонов равна примерно 1 МэВ. Увеличение в 3 раза величины элементарного электрического заряда привело бы к увеличению на порядок постоянной тонкой структуры. Во столько же раз увеличилась бы энергия отталкивания протонов, т. е. ядра гелия стали бы нестабильными

и в мире не смогли бы существовать элементы с  $A \geq 4$ . Трудно представить себе жизнь, основанную лишь на атомах водорода. Не говоря уже о том, что энергия ионизации (постоянная Ридберга) этих атомов увеличилась бы в 81 раз и составила бы (при прежних значениях всех прочих фундаментальных констант) 1,1 МэВ. При таких «ядерных» энергиях можно забыть о понятии валентности, так как никакие химические реакции не будут возможны.

Если бы константа Ферми  $G_F$ , регулирующая интенсивность слабого взаимодействия, была больше, то первые реакции в цепочке термоядерного синтеза, идущие по слабому взаимодействию, протекали бы интенсивнее и звезды вспыхивали и угасали бы гораздо быстрее — их времени жизни могло бы не хватить для зарождения жизни на вращающихся вокруг них планетах. При нынешнем значении  $G_F$  время жизни звезд, подобных Солнцу, составляет около 10 млрд лет, Земля возникла около 5 млрд лет назад, а жизнь на ней появилась примерно 4 млрд лет назад.

Существен и темп расширения Вселенной, зависящий от гравитационной постоянной, плотностей темной энергии, темной материи и вещества в ней, а также от плотности электромагнитного излучения. При высокой суммарной плотности всех видов энергии, когда расширение Вселенной сменяется сжатием, времени жизни Вселенной могло бы не хватить даже на образование звезд, не говоря уже о возникновении жизни на планетах. При низкой плотности энергии темп расширения Вселенной высок, что ведет к неустойчивости галактик. Тонкая настройка параметров Вселенной обеспечила доминирование тех или иных видов энергии на разных этапах эволюции Вселенной, что в конечном итоге и привело к тому миру, который комфортен для нашего в нем проживания.

В обзоре<sup>1</sup> и книге<sup>2</sup> можно найти подробное обсуждение этих вопросов, ряд других интересных примеров тонкой настройки параметров и обширный список литературы. Математически тонкая настройка выражается в соотношениях (обычно неравенствах) между физическими параметрами — в частности, между массами частиц и фундаментальными константами.

Самым удивительным здесь является тот факт, что только тонкая настройка делает возможным существование жизни во Вселенной. Выглядит так, будто Вселенная приспособлена для человека, создана для него. Собственно говоря, именно такой антропоцентристский взгляд веками и тысячелетиями излагался в сказаниях о сотворении мира и человека в разных религиях и учениях древних мыслителей. Но со временем наука отделила природу от человека и стала изучать мир как независимый от людей объект. На этом пути были достигнуты грандиозные успехи, о которых, в сущности, и шла речь в этом курсе. Но в XX в. выяснилось удивительное соответствие значений физических параметров и условий существования живых существ. Вселенную в целом мы не можем рассматривать отстраненно, изучать ее со стороны, как и все другие объекты. Мы сами находимся внутри нее, и это тоже экспериментальный факт. Само наличие наблюдателя уже накладывает сильные ограничения на свойства

<sup>1</sup> И. Л. Розенталь. Физические закономерности и численные значения фундаментальных постоянных. УФН **131**, вып. 2, 1980. С. 239–255

<sup>2</sup> К. А. Томиллин. Фундаментальные физические постоянные в историческом и методологическом аспектах. М.: Физматлит, 2006.

наблюдаемой им Вселенной. Скажем, крокодилы, если бы они умели думать, описывали бы свой мир как теплый и влажный, отлично приспособленный для их существования. А другого мира (Гималаи или тундру) они никогда не видели, потому что крокодилы там не водятся.

Еще в 1957 г. советский астрофизик Г. М. Идлис поставил обсуждаемый вопрос со всей прямотой: «Не являются ли основные черты наблюдаемой астрономической Вселенной простым следствием того, что перед нами не произвольная часть бесконечной в своем многообразии Вселенной, а такая конкретная конечная область ее, в которой жизнь заведомо имела возможность возникнуть и существует в настоящее время налицо? <...> Это очень важное положение, означающее, что мы наблюдаем заведомо не произвольную область Вселенной, а ту, особая структура которой сделала ее пригодной для возникновения и развития жизни» (цит. по книге К. А. Томилина, см. сноску 2).

Эти идеи, развивавшиеся в 1950–1960 гг. целым рядом астрофизиков, вылились в четкую формулировку *антропного принципа*, данную английским математиком и физиком Б. Картером (1970): «То, что мы ожидаем наблюдать, должно быть ограничено условиями, необходимыми для нашего существования как наблюдателей». Иными словами, Вселенная с другими характеристиками принципиально не наблюдаема.

Антропный принцип имеет множество интерпретаций и толкований. Но все согласны, что случайное возникновение нужных значений многих физических параметров при разовом акте творения Вселенной полностью исключено ввиду невообразимо малой вероятности такого совпадения. Остается основной вопрос, на который никто не может однозначно ответить: «Каким образом возникла эта часть Вселенной, в который мы обитаем?»

Один ответ: «Существует множество миров (вселенных), но они для нас ненаблюдаемы, как Гималаи и тундра для крокодилов». Идея множественности миров неоднократно обсуждалась физиками, философами, писателями-фантастами.

Другой ответ: «Вселенная циклически проходит через Большой взрыв, и при многократном повторении этого процесса наблюдаемые нами параметры вполне могут возникнуть и случайным образом». Оба этих ответа остаются на уровне веры, пока не будет создана теория, описывающая появление и свойства ненаблюдаемых вселенных.

Третий ответ: «Появление таких значений физических параметров не случайно: Вселенная и была создана для того, чтобы в ней появились жизнь и разум». Этот ответ целиком основан на вере и вряд ли может быть когда-либо проверен (подтвержден или опровергнут) в экспериментах.

Еще одной крайней точкой зрения является четвертый ответ: «Жизнь есть производная от свойств неживой материи и потому приспособилась именно к таким условиям». Эта точка зрения традиционна для классической науки и имеет в свою поддержку свойства наблюдаемых нами организмов. Человеческий глаз наиболее чувствителен к узкому оптическому диапазону, на который приходится максимум излучения Солнца, — ясно, что это не случайно. Или, например, вес тела пропорционален его массе (т. е. кубу линейного размера) и ускорению свободного падения  $g$ :  $P = agl^3$ . Сила же мышц пропорциональна площади их сечения, т. е. квадрату линейного размера:  $F = bl^2$ . Из неравенства  $F \geq P$

следует, что линейные размеры живых организмов не могут быть слишком велики ( $l \leq b/(ag)$ ), иначе они будут раздавлены собственным весом. На других планетах, с другим значением ускорения свободного падения  $g$ , эти размеры могут быть иными.

В то же время теплокровные живые организмы не могут быть слишком маленькими. Вырабатываемая ими энергия пропорциональна массе тела, т.е. опять-таки кубу линейного размера:  $W_+ = cl^3$ . Тепловое излучение с поверхности тела пропорционально его площади:  $W_- = dl^2$ . Из неравенства  $W_+ \geq W_-$  следует, что размеры тела теплокровных животных не могут быть слишком малы:  $l \geq d/c$ . Простые модели позволяют найти зависимость частоты дыхания и пульса от массы теплокровного организма, а также оценить зависимость от массы отведенного ему природой времени жизни (заметим, что человек выпадает из общей закономерности, в целом подтверждаемой биологами).

Примеров успешного применения физических законов к живым организмам очень много, с ними можно познакомиться в книге<sup>3</sup> и в последней главе предыдущего издания нашего учебника<sup>4</sup>. Но одно дело — приспособиться к условиям существования на отдельно взятой планете и совсем другое — развитие жизни в масштабах всей Вселенной.

Есть, наконец, и такая точка зрения, что антропный принцип ничего не объясняет и представляет собой лишь уклонение от прямого ответа.

Как бы то ни было, мы завершаем наш курс постановкой этой самой, быть может, загадочной проблемы современной науки.

В основном корпусе учебника мы попытались изложить установленные сведения об окружающем нас мире и важнейших закономерностях, действующих в нем, а также старались показать, как достижения фундаментальных наук используются в различных технологиях и устройствах. Но Козьма Прутков прав: «Никто не обнимет необъятного». Главная трудность, которую приходится преодолевать авторам любого учебника, — это отбор наиболее важного материала из всего обширного массива накопленных знаний. Многое оказалось «за кадром», некоторые вопросы пришлось рассмотреть достаточно бегло в расчете на более детальное ознакомление с ними в процессе дальнейшего обучения студентов на специализированных кафедрах.

Авторам остается лишь поблагодарить своих читателей и выразить надежду, что они согласятся с утверждением из старинного учебника физики, изданного в 1787 г. при императрице Екатерине II, который уже цитировался в самом начале нашего курса (см. т. 1, с. 7):

«Большая часть видимого нами мира кажется очам нашим совсем иначе, нежели как она в самом деле находится. Несмотря на сие, рассматривание видимого разнообразия мира имеет также свою пользу и ведет нас к истине».

<sup>3</sup> К. Шмидт-Нильсен. Размеры животных: почему они так важны? М.: Мир, 1987.

<sup>4</sup> Н. П. Калашиков, М. А. Смондырев. Основы физики. Т. 2. М.: Дрофа, 2004.

# Оглавление

<b>Часть IV</b>	<b>Колебания и волны</b>	<b>3</b>
<b>Глава 24</b>	<b>Колебательное движение</b>	<b>5</b>
24.1	Уравнение гармонических колебаний . . . . .	5
	Пружинный маятник . . . . .	5
	Математический маятник . . . . .	6
	Физический маятник . . . . .	7
	Движение поршня в сосуде с идеальным газом . . . . .	8
	Электромагнитный контур . . . . .	8
24.2	Гармонические колебания . . . . .	9
24.3	Сохранение энергии при гармонических колебаниях . . . . .	13
24.4	Сложение однонаправленных колебаний . . . . .	15
	Сложение колебаний с одинаковыми частотами . . . . .	15
	Биения . . . . .	16
	Колебания двух связанных осцилляторов . . . . .	18
24.5	Сложение взаимно перпендикулярных колебаний . . . . .	21
24.6	Свободные затухающие колебания . . . . .	23
	Уравнение затухающих колебаний . . . . .	23
	Анализ решений . . . . .	24
24.7	Вынужденные колебания . . . . .	28
	Контрольные вопросы . . . . .	31
<b>Глава 25</b>	<b>Переменный ток</b>	<b>33</b>
25.1	Квазистационарные токи . . . . .	33
25.2	Переменный ток через элементы цепи . . . . .	33
	Переменный ток через сопротивление . . . . .	33
	Переменный ток через индуктивность . . . . .	34
	Переменный ток через емкость . . . . .	35
25.3	Цепь переменного тока . . . . .	35
25.4	Резонансные явления . . . . .	37
	Резонанс напряжений . . . . .	38
	Резонанс токов . . . . .	38
25.5	Мощность в цепи переменного тока . . . . .	41
	Контрольные вопросы . . . . .	42



<b>Глава 26</b>	<b>Нелинейные колебания</b>	<b>43</b>
26.1	Нелинейные колебания маятника . . . . .	44
26.2	Фазовый портрет . . . . .	47
26.3	Автоколебания . . . . .	50
26.4	Эволюция и взаимодействие популяций . . . . .	54
	Популяция в отсутствие сдерживающих факторов . . . . .	54
	Внутривидовая конкуренция . . . . .	55
	О взаимоотношениях зайцев и волков . . . . .	57
	Контрольные вопросы . . . . .	60
<b>Глава 27</b>	<b>Волновые процессы</b>	<b>61</b>
27.1	Волны в упругих средах . . . . .	61
	Колебания струны . . . . .	61
	Колебания в идеальном газе . . . . .	62
	Колебания в твердых телах . . . . .	64
27.2	Решение волнового уравнения . . . . .	64
27.3	Энергия волны . . . . .	69
	Применения к звуковой волне . . . . .	70
27.4	Стоячие волны . . . . .	72
	Струна, закрепленная на одном конце . . . . .	73
	Гармоники . . . . .	74
	Сложение гармоник . . . . .	75
	Спектр колебаний . . . . .	78
27.5	Сферические волны . . . . .	80
	Трехмерное волновое уравнение . . . . .	81
	Эффект Доплера для звуковых волн . . . . .	82
	Сверхзвуковые скорости . . . . .	87
	Контрольные вопросы . . . . .	90
<b>Часть V</b>	<b>Основы волновой оптики</b>	<b>93</b>
<b>Глава 28</b>	<b>Волновая теория света и законы геометрической оптики</b>	<b>95</b>
28.1	Принцип Гюйгенса . . . . .	95
28.2	Законы отражения и преломления света . . . . .	96
	Закон отражения . . . . .	96
	Закон преломления . . . . .	97
	Принцип Ферма . . . . .	100
	Контрольные вопросы . . . . .	101
<b>Глава 29</b>	<b>Интерференция света</b>	<b>103</b>
29.1	Интенсивность света . . . . .	103
29.2	Когерентность световых волн . . . . .	104
29.3	Интерференция света от двух источников . . . . .	107
29.4	Способы наблюдения интерференции света . . . . .	109

29.5	Интерференция света в тонких пленках . . . . .	110
29.6	Полосы равной толщины. Кольца Ньютона . . . . .	112
29.7	Интерферометры . . . . .	114
	Контрольные вопросы . . . . .	116
<b>Глава 30</b>	<b>Дифракция света</b> . . . . .	<b>117</b>
30.1	Принцип Гюйгенса—Френеля . . . . .	117
30.2	Метод зон Френеля . . . . .	119
30.3	Дифракция на круглом отверстии и диске . . . . .	122
30.4	Дифракция Фраунгофера от щели . . . . .	124
30.5	Дифракционная решетка . . . . .	129
	Дифракция от двух щелей . . . . .	130
	Дифракция на решетке . . . . .	131
	Характеристики дифракционной решетки . . . . .	134
30.6	Дифракция рентгеновских лучей . . . . .	136
30.7	Голография . . . . .	138
	Контрольные вопросы . . . . .	141
<b>Глава 31</b>	<b>Влияние среды на свойства света</b> . . . . .	<b>143</b>
31.1	Поляризация света . . . . .	143
	Поляризация электромагнитной волны . . . . .	143
	Естественный и поляризованный свет . . . . .	144
	Закон Малюса . . . . .	146
	Поляризация при отражении и преломлении . . . . .	148
31.2	Дисперсия света . . . . .	150
	Теория радуги . . . . .	153
	Элементарная теория дисперсии . . . . .	157
31.3	Поглощение света . . . . .	160
31.4	Рассеяние света . . . . .	161
	Контрольные вопросы . . . . .	161
<b>Глава 32</b>	<b>Электромагнитные волны</b> . . . . .	<b>163</b>
32.1	Усреднение микроскопических полей . . . . .	163
32.2	Уравнения Максвелла для сплошных сред . . . . .	166
32.3	Волновое уравнение . . . . .	168
32.4	Основные свойства электромагнитных волн . . . . .	169
32.5	Эффект Доплера для электромагнитных волн . . . . .	172
	Отражение света от движущегося зеркала . . . . .	174
32.6	Энергия и импульс электромагнитного поля . . . . .	176
	Давление света . . . . .	178
	Солнечный парус . . . . .	179
32.7	Законы оптики и уравнения Максвелла. Формулы Френеля . . . . .	180
	Отражение и преломление $s$ -поляризованной волны . . . . .	181
	Отражение и преломление $p$ -поляризованной волны . . . . .	184

32.8	Полное внутреннее отражение . . . . .	186
32.9	Затухание волн в металле. Скин-эффект . . . . .	188
32.10	Дипольное излучение . . . . .	191
32.11	Излучение равномерно движущегося заряда . . . . .	197
	Контрольные вопросы . . . . .	199
<b>Часть VI Основы квантовой механики</b>		<b>201</b>
<b>Глава 33</b>	<b>Квантовая природа излучения</b>	<b>203</b>
33.1	Тепловое излучение . . . . .	203
33.2	Эмпирические законы излучения абсолютно черного тела . . . . .	209
33.3	Классические результаты для теплового излучения . . . . .	211
33.4	Закон излучения Планка . . . . .	213
	Контрольные вопросы . . . . .	218
<b>Глава 34</b>	<b>Фотоны</b>	<b>219</b>
34.1	Кванты света . . . . .	219
34.2	Фотоэлектрический эффект . . . . .	221
34.3	Эффект Комптона . . . . .	225
34.4	Опыт Боте . . . . .	230
	Контрольные вопросы . . . . .	232
<b>Глава 35</b>	<b>Волновые свойства микрочастиц</b>	<b>233</b>
35.1	Атом Бора . . . . .	233
	Постулаты Бора . . . . .	234
	Характеристики водородоподобного атома . . . . .	235
35.2	Волны материи . . . . .	241
35.3	Соотношения неопределенностей . . . . .	246
	Некоторые следствия соотношений неопределенностей . . . . .	247
35.4	Двухщелевой эксперимент . . . . .	251
35.5	О границах применимости квантовой механики . . . . .	252
	Контрольные вопросы . . . . .	255
<b>Глава 36</b>	<b>Уравнение Шрёдингера</b>	<b>257</b>
36.1	Волна вероятности . . . . .	257
36.2	Общее уравнение Шрёдингера . . . . .	259
36.3	Операторы, симметрия и законы сохранения . . . . .	261
36.4	Стационарное уравнение Шрёдингера . . . . .	262
36.5	Уравнение Шрёдингера для простейших систем . . . . .	265
	Свободная частица, движущаяся вдоль оси $x$ . . . . .	265
	Частица в бесконечно глубокой потенциальной яме . . . . .	266
	Частица в трехмерной потенциальной яме . . . . .	268
	Одномерный осциллятор . . . . .	269
	Трехмерный осциллятор . . . . .	270

36.6	Принцип соответствия Бора . . . . .	271
36.7	Отражение и туннелирование частиц . . . . .	273
	Низкий бесконечный барьер . . . . .	273
	Высокий бесконечный барьер . . . . .	275
	Конечный потенциальный барьер . . . . .	276
36.8	Оптическая аналогия прохождения частицы над барьером . . . . .	281
	Ступенчатый потенциал . . . . .	281
	Прямоугольный барьер конечной ширины . . . . .	283
	Прохождение света через многослойную структуру . . . . .	285
	Контрольные вопросы . . . . .	287
<b>Глава 37</b>	<b>Теория атома</b>	<b>289</b>
37.1	Коммутирующие операторы . . . . .	289
37.2	Момент количества движения . . . . .	291
37.3	Атом водорода . . . . .	295
37.4	Спин электрона и тонкая структура спектров . . . . .	303
	Опыт Штерна—Герлаха . . . . .	306
37.5	Векторная модель атома . . . . .	310
37.6	Принцип Паули и валентность элементов . . . . .	312
	Эффективный заряд ядра, оценки потенциала ионизации и за- кон Мозли . . . . .	314
	Электронная конфигурация атомов . . . . .	316
37.7	Принцип тождественности частиц . . . . .	317
	Контрольные вопросы . . . . .	320
<b>Глава 38</b>	<b>Физическая природа химической связи</b>	<b>321</b>
38.1	Молекулы . . . . .	321
	Ионная связь . . . . .	321
	Ковалентная связь . . . . .	323
	Комбинации различных типов связи . . . . .	328
38.2	Пространственное строение молекул . . . . .	330
38.3	Молекулярные спектры . . . . .	332
	Электронные уровни энергии . . . . .	333
	Энергия колебательного движения ядер . . . . .	333
	Энергия вращательного движения молекул . . . . .	334
	Молекулярные спектры . . . . .	335
38.4	Комбинационное рассеяние света . . . . .	336
<b>Часть VII</b>	<b>Основы теории строения вещества</b>	<b>339</b>
<b>Глава 39</b>	<b>Взаимодействие излучения с веществом</b>	<b>341</b>
39.1	Вывод формулы Планка по Эйнштейну . . . . .	341
39.2	Общие сведения о лазерах . . . . .	345
	Прохождение излучения через вещество. Инверсия населеннос- тей . . . . .	345

	Принцип работы лазера . . . . .	346
	Типы лазеров . . . . .	349
39.3	Нелинейные эффекты в оптике . . . . .	352
	Контрольные вопросы . . . . .	355
<b>Глава 40</b>	<b>Теплоемкость кристаллов. Статистика Бозе—Эйнштейна</b>	<b>357</b>
40.1	Классические представления о теплоемкости кристаллов . . . . .	357
40.2	Средняя энергия квантового осциллятора . . . . .	359
	Средняя энергия квантового ротатора . . . . .	361
40.3	Теплоемкость кристаллов по Эйнштейну . . . . .	363
40.4	Теория теплоемкости твердых тел Дебая . . . . .	364
	Число колебаний в единице объема . . . . .	364
	Характеристическая температура Дебая . . . . .	366
40.5	Фононы . . . . .	368
	Линейная цепочка одинаковых атомов . . . . .	369
	Линейная двухатомная цепочка . . . . .	370
	Квантование колебаний кристаллической решетки . . . . .	373
40.6	Статистика Бозе—Эйнштейна . . . . .	375
	Плотность квантовых состояний . . . . .	378
	Конденсация Бозе—Эйнштейна . . . . .	379
	Контрольные вопросы . . . . .	382
<b>Глава 41</b>	<b>Элементы физики твердого тела. Статистика Ферми—Дирака</b>	<b>383</b>
41.1	Энергетические зоны в твердых телах . . . . .	383
	Применение оптической аналогии . . . . .	383
	Формирование энергетических зон в кристалле . . . . .	384
41.2	Металлы, диэлектрики и полупроводники . . . . .	388
41.3	Теория свободных электронов в металле. Энергия Ферми . . . . .	390
	Электронный газ при нулевой температуре . . . . .	393
41.4	Статистика Ферми—Дирака . . . . .	396
41.5	Распределение Ферми—Дирака . . . . .	396
41.6	Эффективная масса электрона . . . . .	404
41.7	Электропроводность металлов . . . . .	408
41.8	Полупроводники . . . . .	409
	Дырки . . . . .	409
	Собственная проводимость полупроводников . . . . .	412
	Примесная проводимость полупроводников . . . . .	415
	Контрольные вопросы . . . . .	420
<b>Глава 42</b>	<b>Контактные явления</b>	<b>421</b>
42.1	Работа выхода и контактная разность потенциалов в металле . . . . .	421
42.2	Термоэлектрические явления . . . . .	424
	Термо-ЭДС . . . . .	425

	ЭДС термопары . . . . .	428
	Эффект Пельтье . . . . .	431
	Эффект Томсона . . . . .	434
42.3	Полупроводниковые выпрямители . . . . .	438
	Контактная разность потенциалов в $p$ - $n$ -переходе . . . . .	438
	Сопротивление и односторонняя проводимость $p$ - $n$ -перехода . . . . .	439
	Туннельный диод . . . . .	441
	Фотоэлектрические явления в полупроводниках . . . . .	444
	Контрольные вопросы . . . . .	448
<b>Глава 43</b>	<b>Макроскопические квантовые явления</b>	<b>449</b>
43.1	Сверхтекучесть . . . . .	449
	Элементы микроскопической теории сверхтекучести . . . . .	450
	Сверхтекучесть конденсата . . . . .	452
43.2	Сверхпроводимость . . . . .	453
	Эффект Мейснера и критические значения температуры и магнитного поля . . . . .	454
	Сверхпроводники 2-го рода и вихри Абрикосова . . . . .	457
	Промежуточное состояние . . . . .	461
	Уравнение Лондонов . . . . .	462
	Квантование магнитного потока . . . . .	464
	Электрон-фононное взаимодействие и куперовские пары . . . . .	466
	Микроскопическая теория сверхпроводимости (БКШ) . . . . .	467
	Туннелирование электронов при контакте сверхпроводников . . . . .	471
	Эффекты Джозефсона . . . . .	474
	Контрольные вопросы . . . . .	481
<b>Часть VIII</b>	<b>Основы физики микромира</b>	<b>483</b>
<b>Глава 44</b>	<b>Физика атомного ядра</b>	<b>485</b>
44.1	Состав атомного ядра . . . . .	485
44.2	Физическая природа ядерных сил . . . . .	487
44.3	Масса и дефект массы ядра . . . . .	489
44.4	Модели атомного ядра . . . . .	493
	Модель жидкой капли . . . . .	493
	Оболочечная модель ядра . . . . .	496
44.5	Радиоактивность . . . . .	499
	Закон радиоактивного распада . . . . .	500
	Распад «дочерних ядер» . . . . .	502
	Альфа-распад . . . . .	505
	Бета-распад . . . . .	508
	Ядерные реакции и определение возраста археологических образцов радиоуглеродным методом . . . . .	510
44.6	Элементы дозиметрии . . . . .	518
44.7	Атомная и термоядерная энергетика . . . . .	524

Деление ядер . . . . .	524
Цепная реакция. Атомные бомбы и реакторы . . . . .	528
Термоядерная энергетика . . . . .	538
Контрольные вопросы . . . . .	543
<b>Глава 45 Основные представления физики элементарных частиц</b>	<b>545</b>
45.1 Фундаментальные взаимодействия . . . . .	545
Квантовая электродинамика . . . . .	546
Сильные ядерные взаимодействия . . . . .	549
Слабые взаимодействия . . . . .	553
Нарушение $P$ -, $C$ - и $CP$ -симметрии в слабых взаимодействиях	557
45.2 Квантовая хромодинамика . . . . .	559
45.3 Кварки и лептоны в Стандартной теории . . . . .	563
Контрольные вопросы . . . . .	566
<b>Глава 46 Элементарные частицы и космология</b>	<b>567</b>
46.1 За пределами Стандартной модели . . . . .	568
Масса нейтрино . . . . .	568
Барионная асимметрия Вселенной . . . . .	569
Великое объединение . . . . .	570
Суперсимметрия . . . . .	571
Суперструны . . . . .	572
46.2 Вселенная — прошлое и будущее . . . . .	573
Расширяющаяся Вселенная и закон Хаббла . . . . .	573
Критическая плотность . . . . .	580
Темная материя . . . . .	581
Темная энергия . . . . .	585
Гравитация и планковские масштабы . . . . .	588
Горячая Вселенная и Большой взрыв . . . . .	590
Контрольные вопросы . . . . .	594
<b>ЭПИЛОГ</b>	<b>595</b>

*Минимальные системные требования определяются соответствующими требованиями программ Adobe Reader версии не ниже 11-й либо Adobe Digital Editions версии не ниже 4.5 для платформ Windows, Mac OS, Android и iOS; экран 10"*

*Учебное электронное издание*

Серия: «Учебник для высшей школы»

**Калашников Николай Павлович**  
**Смондырев Михаил Александрович**

## **ОСНОВЫ ФИЗИКИ**

В трех томах

Том 2

Ведущие редакторы *И. Я. Иццоки, Т. Г. Хохлова*

Художники *В. Е. Шкерин, В. А. Прокудин*

Корректор *Н. В. Бурдина*

Оригинал-макет подготовлен *М. А. Смондыревым* в пакете  $\text{\LaTeX}$  2 $\epsilon$

Подписано к использованию 20.04.21.

Формат 150×225 мм

Издательство «Лаборатория знаний»

125167, Москва, проезд Аэропорта, д. 3

Телефон: (499) 157-5272

e-mail: [info@pilotLZ.ru](mailto:info@pilotLZ.ru), <http://www.pilotLZ.ru>

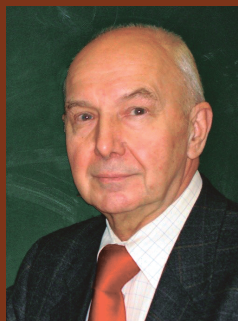


Данный учебник завершает двухтомное издание «Основы физики» и входит в единый комплект вместе со сборником задач и упражнений тех же авторов. Предыдущее издание учебника в составе других учебных материалов выдвинуто на соискание премии Правительства Российской Федерации в области образования. Настоящее издание переработано и существенно дополнено.

Во втором томе, помимо традиционных разделов, включающих колебания и волны, основы квантовой механики, строения вещества и ядерной физики, представлены многие другие интересные темы, которые обычно отсутствуют в учебниках: нелинейные колебания и колебания в живой природе, макроскопические квантовые явления, сверхтекучесть и сверхпроводимость, контактные явления (в том числе контакты сверхпроводников), основы физики элементарных частиц и последние достижения космологии.

Учебник соответствует программе дисциплины «Физика» для естественнонаучных и технических университетов и не имеет аналогов по фундаментальности подхода, с одной стороны, и по живости и доступности изложения – с другой.

**КАЛАШНИКОВ Николай Павлович.** Профессор, доктор физико-математических наук, заслуженный деятель науки РФ, заведующий кафедрой общей физики Национального исследовательского ядерного университета «МИФИ» (НИЯУ МИФИ).



Автор более 380 научных трудов и 14 монографий.

С 2012 г. директор агентства по аккредитации образовательных программ инженерных специальностей.

Основные направления научных исследований профессора Калашникова – ядерная физика, взаимодействия быстрых заряженных частиц с веществом, физика твердого тела в экстремальных состояниях.

**СМОНДЫРЕВ Михаил Александрович.** Профессор, доктор физико-математических наук, ведущий научный сотрудник Лаборатории теоретической физики им. Н. Н. Боголюбова Объединенного института ядерных исследований (ОИЯИ, Дубна), профессор Московского авиационного института (Национального исследовательского университета).



В сферу научных интересов входят вопросы теории элементарных частиц и твердого тела. Автор и переводчик целого ряда научно-популярных книг и статей.

Лауреат премии Ленинского комсомола и премий ОИЯИ – Международной межправительственной научно-исследовательской организации.